131. Pyranosyl-RNA: Further Observations on Replication[1])

by **Martin Bolli, Ronald Micura, Stefan Pitsch,** and **Albert Eschenmoser***

Laboratory of Organic Chemistry, Swiss Federal Institute of Technology, ETH-Zentrum, Universitätstrasse 16, CH-8092 Zürich

and

The SKAGGS Institute for Chemical Biology at *SCRIPPS* Research Institute, 10550 North Torrey Pines Road, La Jolla/CA 92037, USA

Dedicated to *Dieter Seebach* on the occasion of his 60th birthday

(23.VI.97)

Replication (single-turnover) of pyranosyl-RNA (= p-RNA) sequences can be accomplished reliably by template-directed ligation of 2',3'-cyclophosphates of short oligomers. The ligation process was studied using (mostly) octamers as templates and tetramers as ligands. The transcription of the sequence pr(GGGCGGGC) into the (antiparallel) complementary sequence pr(GCCCGCCC) by ligation of two molecules of pr(GCCC)-2',3'-cp was investigated in detail: In aqueous 1.5M LiCl solution of pH 8.5 at room temperature (0.45 mM ligand, 0.15 mM template), the reaction proceeds in up to 60% yield within a week. It is limited by concomitant hydrolysis of cyclophosphate groups of both reactand and ligation product as the only efficient side reaction, the latter occurring *ca.* three times more slowly than ligation. No ligation at all is observed in the absence of template. The reaction is highly regioselective: the (4' → 2') phosphodiester junction is formed exclusively; no isomeric (4' → 3') junctions are found. For ligation to occur, template and ligand must be homochiral and must have the same sense of chirality; with chiro-diastereoisomeric tetramer-2',3'-cyclophosphates containing a single enantio-ribopyranosyl unit, no ligation is observed, except to a minor extent in the case of the diastereoisomer that has that unit at the 4'-end. Observations made in experiments involving six different octamer templates containing isomeric base sequences indicate that the ligation process does not tolerate a mismatch at ligation sites. However, ligation still takes place when a mismatch occurs at either end of the (octamer) template. Ligation efficiencies differ widely, depending on the nature, as well as the sequence, of participating bases. These differences can be understood qualitatively by considering the relative stability of ternary pre-ligation complexes, together with the differences in *inter*strand base stacking at ligation sites. Dominance of the latter over *intra*strand base stacking is the feature of the p-RNA structure that appears to determine most of the characteristic properties of p-RNA.

As regards the etiological context of our work on nucleic-acid alternatives, it is essential that the chemical properties found for p-RNA be compared with the corresponding properties of RNA. In the RNA series, the two ligations of the replicative cycle r(GGGCGGGC) ⇔ r(GCCCGCCC) using the corresponding ribofuranosyl-tetramer 2',3'-cyclophosphates as ligands are found to proceed also, though somewhat less efficiently than in the p-RNA series; however, the ligation step produces exclusively the unnatural (5' → 2') phosphodiester junctions instead of the natural (5' → 3') junctions. This is in sharp contrast to p-RNA, where template-controlled 2',3'-cyclophosphate ligations produce the 'correct' phosphodiester junctions.

1. Introduction. – In the ETH laboratory, recent work on the chemistry of pyranosyl-RNA (= p-RNA) has focused on the topics of template-directed copying of base se-

[1]) 5th Communication on the chemistry of p-RNA. 4th Communication: [1]. The paper is also communication No. 20 in the series '*Chemistry of α-Amino Nitriles*'; [1] counts as No. 19 in this series, for No. 17 and 18, see [2] and [3], respectively

quences (ligative replication) [4], heterochiral base-pairing [3], hairpin formation [5], and self-assembly of longer base sequences by chiroselective oligomerization of tetramer-2′,3′-cyclophosphates [6]. Moreover, *Jaun* and *Schlönvogt* [1] have recently accomplished a structure determination of a p-RNA duplex by NMR spectroscopy, complemented by a molecular-dynamics study of *Wolf* [1]. Here, we summarize observations of our studies on copying p-RNA base sequences by template-directed ligation of tetramer-2′,3′-cyclophosphates, a process on which we reported our first observations two years ago [4] (*Fig. 1*). As discussed earlier [4–7], the copying of p-RNA sequences is studied in the context of a chemical etiology of nucleic-acid structure. However, the process also deserves attention as a purely chemical one, focusing on molecular recognition within supramolecular assemblies. Template-directed syntheses of oligonucleotide sequences in general [8–16], and of p-RNA sequences in particular [4], exemplify a type of chemo-, regio-, and diastereoselective reactivity, the selectivity of which is brought about by favorable constellational positioning of reaction partners through noncovalent bonding within a supramolecular complex. The potentials such a process possesses, and the problems it poses, are prototypical for the type of chemistry ('informational chemistry'), of which coded intermolecular transfer of combinatorial structural information is an example.

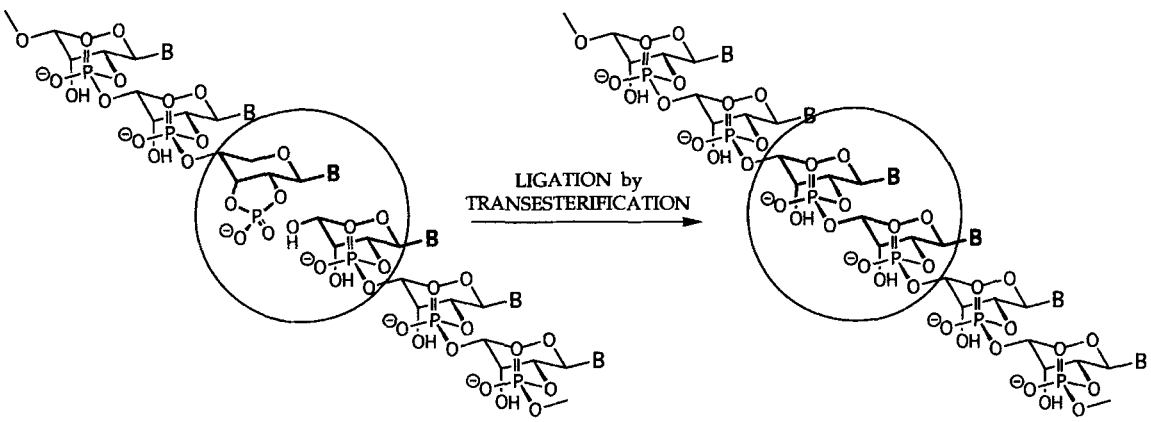


Fig. 1. *Formation of p-RNA strands by (4′ → 2′) ligation of short oligomers using 2′,3′-cyclophosphate groups as activated forms of 2′-phosphate groups*

A systematic study of template-controlled replication in the p-RNA series must address problems such as the constitutional nature of phosphate activation, the dependence of the ligation rate as well as of the ligation/hydrolysis ratio on reaction conditions, the regioselectivity, fidelity and stereoselectivity of ligation, the dependence of these properties on the template sequence as well as on the sequence length of the ligands, and, finally, the problem of overcoming or circumventing product inhibition. In view of the etiological context within which our work on p-RNA is carried out, it is also necessary to compare p-RNA and RNA with respect to their performance in template-directed

replicative ligation under similar reaction conditions with the same type of phosphate activation. In this progress report, we describe observations made in experiments that refer to most of the topics listed above. All experiments described in this paper were carried out with homochiral oligonucleotides of the D-series, except when stated otherwise.

2. Phosphate Activation. – As described in our previous communication [4], we originally used the water-soluble carbodiimide derivative DEC[2]), a reagent well-known to induce template-directed ligation in the DNA series [9] [13] [14], to achieve ligation of tetramer-2′-phosphates on octamer templates. Occurrence of side reactions between (especially unpaired) guanine residues and DEC during the long reaction time required for ligation (weeks), as well as partial transformation of ligand- and product-2′-phosphates into corresponding 2′,3′-cyclophosphates under the reaction conditions (1.5M NaCl, 0.125M $MgCl_2$, 0.1M HEPES buffer, pH 6.3; 40°; large excess of DEC) led us to search for conditions under which the mild degree of phosphate activation intrinsic to the 2′,3′-cyclophosphate group would suffice for template-mediated intermolecular phosphodiester formation to occur. After conditions described in the literature [11] [12][3]) for template-directed ligation of ribofuranosyl-2′,3′-cyclophosphates turned out to be unsuccessful in exploratory experiments in the ribopyranosyl series, a systematic search for conditions led to a protocol with which ligative replication of the guanine-rich octamer sequence pr(4′GGGCGGGC2′) could be reliably accomplished (1.0M NaCl, 0.025M $MgCl_2$, 0.1M HEPES, pH 8.0, room temperature, weeks) [4]. The results were encouraging in the sense that ligation was found to proceed in the presence of the template only, and that yield limitation appeared to be due not to any irreversible side reaction, but exclusively to the (potentially reversible) hydrolytic deactivation of cyclophosphate end groups.

All work described in this report refers to template-directed ligation with ribopyranosyl-2′,3′-cyclophosphates as the activated form of 2′-phosphate end groups (*Fig. 1*). This ligation mode brings about phosphodiester formation in the (2′ → 4′) strand direction, corresponding to the (3′ → 5′) direction in the RNA series and, therefore, opposite in direction to that of strand growth in DNA replication and transcription. In the p-RNA series, cyclophosphate-mediated phosphodiester formation in the 'natural' direction would be equally conceivable and would require ribopyranosyl-3′,4′-cyclophosphate end groups. Our initial attempts to observe ligation in the p-RNA series were actually pursued not with 2′-phosphates and 2′,3′-cyclophosphates, but with the (more easily synthesized) 4′-phosphates and 3′,4′-cyclophosphates of small oligonucleotide units (including pr(C_4)-4′-phosphate and pr(G_4)-4′-phosphate with templates pr(G_8) and pr(C_8), respectively). Since we did not observe ligation under conditions which were found later to be successful with the corresponding 2′-phosphates and 2′,3′-cyclophosphates[4]),

---

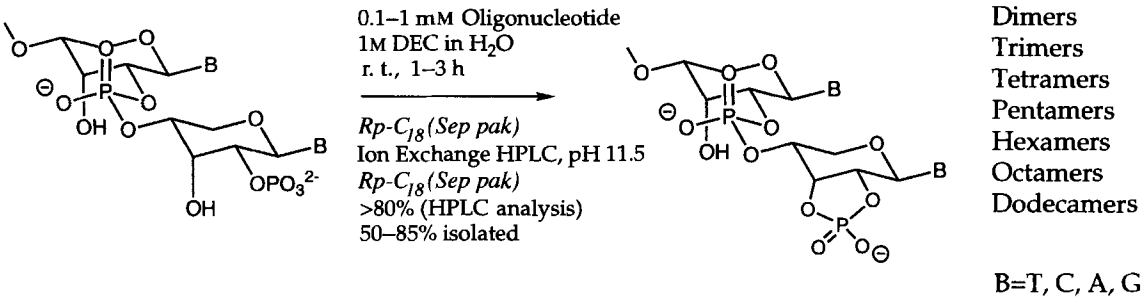
[2]) *N*-[3-(Dimethylamino)propyl]-*N*′-ethylcarbodiimide [9].

[3]) For an earlier investigation on ribofuranosyl-2′,3′-cyclophosphates by *Orgel* and co-workers [10] [11], see the discussion below in *Chapt. 9*.

[4]) Neither pr(C_4)-3′,4′-cyclophosphate nor pr(G_4)-3′,4′-cyclophosphate ligate in the presence of the corresponding octamer templates under conditions where the 2′,3′-cyclophosphates were later found to ligate smoothly. Model considerations at the qualitative level (see discussion of *Figs. 10* and *12*) make this difference understandable.

no further investigation on 3′,4′-cyclophosphate ligation has been carried out to date.

2′,3′-Cyclophosphate derivatives of unprotected oligonucleotides containing any of the nucleobases A, G, C, or T can be prepared by *short* treatment of the corresponding 2′-phosphate derivatives with DEC in aqueous solution at room temperature; yields are nearly quantitative according to HPLC analysis (pH 11.5) and (with some exceptions) above 60% after isolation (*Fig. 2*) [17]. The oligonucleotide (mostly tetramer)-2′-phosphate derivatives used were obtained by automated solid support synthesis [18] using a solid support starter group of the type described in [19][5][6]. 2′,3′-Cyclophosphate derivatives in the p-RNA series can be analyzed by HPLC at pH 11.5 (see caption of *Fig. 4* for conditions) without discernable hydrolytic decomposition.



DEC = *N*-[3-(dimethylamino)propyl]-*N*′-ethylcarbodiimide

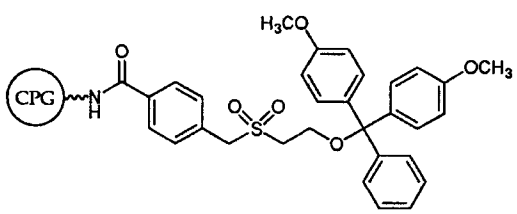
Fig. 2. *Formation of 2′,3′-cyclophosphates from 2′-phosphates of p-RNA monomers or oligomers by dehydration with water-soluble carbodiimide DEC* (for details, see [17])

3. Reaction Conditions. – As described in [4], and in further detail in the present study (*Figs. 3–7*), the transcription of the sequence pr(GGGCGGGC) into pr(GCCCGCCC) has served as the exemplary transformation in our search for optimized reaction conditions for ligative replication. *Table 1* summarizes the course and outcome of the search.

The previously reported conditions [4] (*Entry 1*, *Table 1*) had the weakness of an unfavorable rate ratio between ligation and cyclophosphate-group hydrolysis (*ca.* 1:2). Such a ratio constitutes an effective barrier to the attainment of high transcription yields (see Fig. 7C, a–c, in [4]). Probing this ratio by substituting $CaCl_2$ for $MgCl_2$ (*Entry 2*) had essentially no effect, apart from leading to discernable template decomposition.

---

[5]) The formula of this material is:



[6]) We thank Dr. *H. Moser*, Dr. *R. Häner*, and Dr. *F. Natt* (*Ciba-Geigy AG*, Basel) for a generous gift of this material.

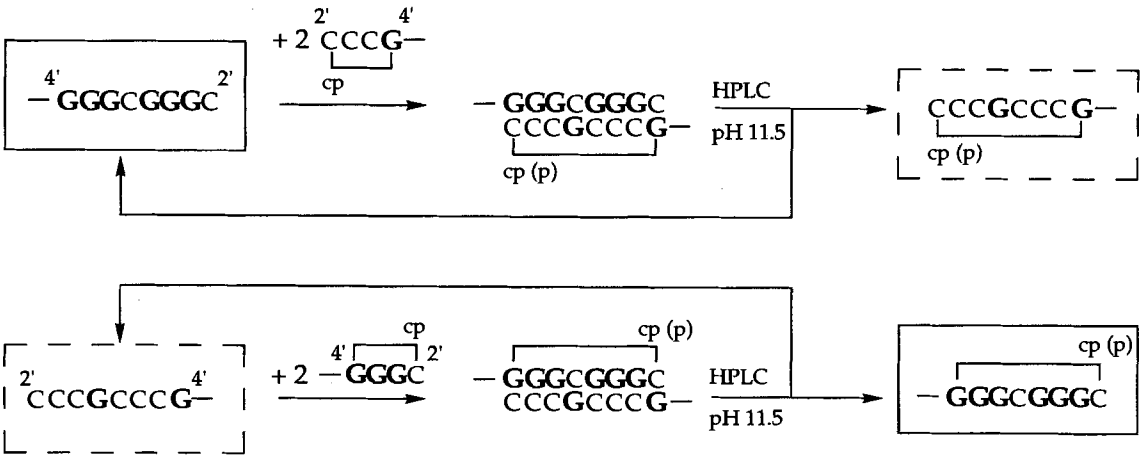


Fig. 3. *Ligative replication of the guanine-rich p-RNA base sequence pr(GGGCGGGC).* For reaction conditions and product yields, see *Table 1* and *Figs. 4–7*

$MnCl_2$ (*Entry 3*), on the other hand, shifted the ratio to the worse, strongly accelerating cyclophosphate hydrolysis as well as strand scission. The latter can be suppressed by lowering $MnCl_2$ concentration by a factor of ten, but cyclophosphate hydrolysis still dominates ligation. An experiment *without* any divalent metal salt (*Entry 5*) led to the surprising and important result that ligation still proceeds, and, although very slow, it is also very clean, shows a more favorable ligation/hydrolysis ratio (2:1), and, therefore, gives better product yields (*Fig. 4, a*). This welcome observation induced a series of experiments in the absence of divalent cations, to explore the effect of the monovalent cation of the bulk electrolyte of the reaction medium. The experiment with 1M KCl (*Entry 6*) pointed to slower ligation, a trend that was interpreted to be related to lower pre-ligation duplex stability. In an experiment with Me_4NCl as the bulk electrolyte (*Entry 7*), almost no ligation was observed. Not unexpectedly then, ligation turned out to be a high-yield reaction in LiCl solution (*Entry 9*; *Fig. 4, b*), with the important advantage of also being markedly faster than the corresponding reaction with NaCl. Exploring the LiCl trail (*Entries 8* and *10–13*) uncovered that raising the pH from 7.0 to 9.0 accelerates ligation and improves the ligation/cyclophosphate hydrolysis ratio from *ca.* 1:2 to 3:1. Doubling the LiCl concentration to 2M, while keeping the pH value of 8.0, had about the same effect as raising the pH value to 9.0; however, the product yield became limited, due to increasingly significant strand scission[7]). As a compromise between yield and reaction rate, the parameters 1.5M LiCl and pH 8.5 (*Entry 13*; *Figs. 4, c*, *5*, and *6*) were chosen for the systematic exploration of selectivity in the ligation process (= 'standard conditions', including 1 mM EDTA, 100 mM HEPES, room temperature)[8]).

---

[7]) Strand scission in template and ligand became the overriding process, when the LiCl concentration was 6M (room temperature).

[8]) The ligation/hydrolysis ratio was found to be unchanged in H_2O/MeCN 2:1 (1M LiCl, pH 8.0). Both ligation and hydrolysis were much slower in H_2O/DMF 3:7 (1M LiCl, pH 8.0).

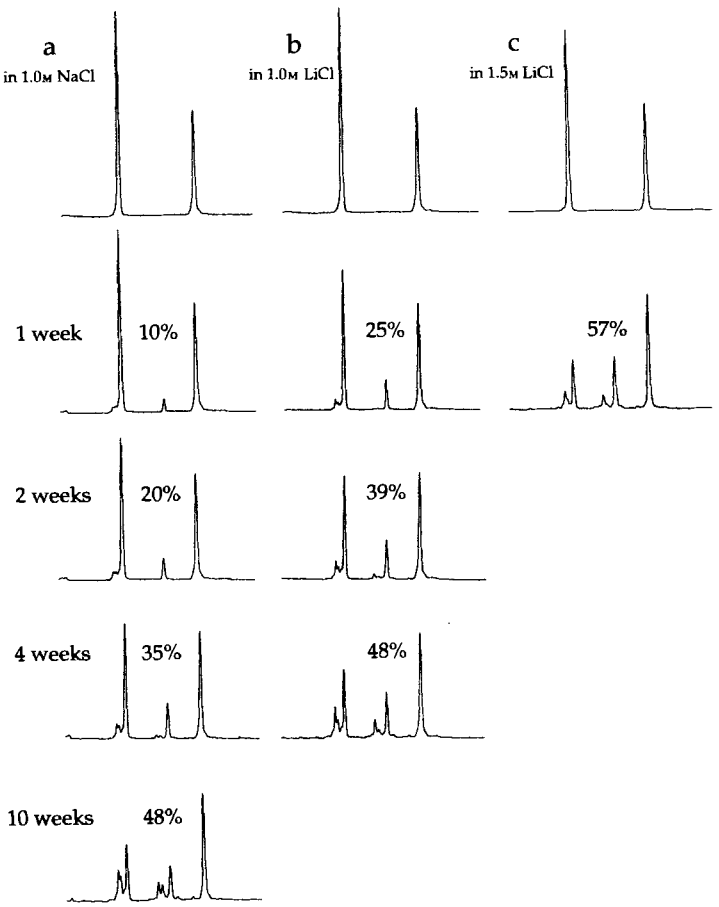


Fig. 4. *HPLC Traces of the ligation of 450* μM *pr(GCCC)-2',3'-cp in the presence of 150* μM *pr(GGGCGGGC) as template at room temperature in* a) *1*M *NaCl, 5* mM *EDTA, 100* mM *HEPES, pH 8.0,* H_2O; b) *1*M *LiCl, 1* mM *EDTA, 100* mM *HEPES, pH 8.0,* H_2O; c) 1.5M *LiCl, 1* mM *EDTA, 100* mM *HEPES, pH 8.5,* H_2O. Reaction times and yields of ligated products are given. Traces at the top refer to the start of the reaction. For peak assignments, see *Fig. 5*. HPLC: *Mono Q HR 5/5*, elution with 10 mM sodium phosphate buffer, pH 11.5, with a linear gradient of 0–1M NaCl within 30 min, detection at 270 nm.

Fig. 7 shows yield *vs.* reaction-time plots of the replicative copying of the two sequences pr(GGGCGGGC) and pr(GCCCGCCC). For product characterization, see *Chapt. 4*.

In each case that reference is made to cyclophosphate hydrolysis in ligands or ligation products, opening of the cyclophosphate ring to a mixture of corresponding 2'- and 3'-phosphates is implied. The 2'-isomer is easily identified by HPLC (since 2'-phosphates serve as starting materials for the preparation of the cyclophosphates) and can be converted to the 2',3'-cyclophosphate by treatment with DEC under mild conditions (see *Chapt. 2*). It represents the main component of the hydrolysis mixture in all known cases in a ratio that depends on the nature of the neighboring nucleobase. The 3'-phosphates

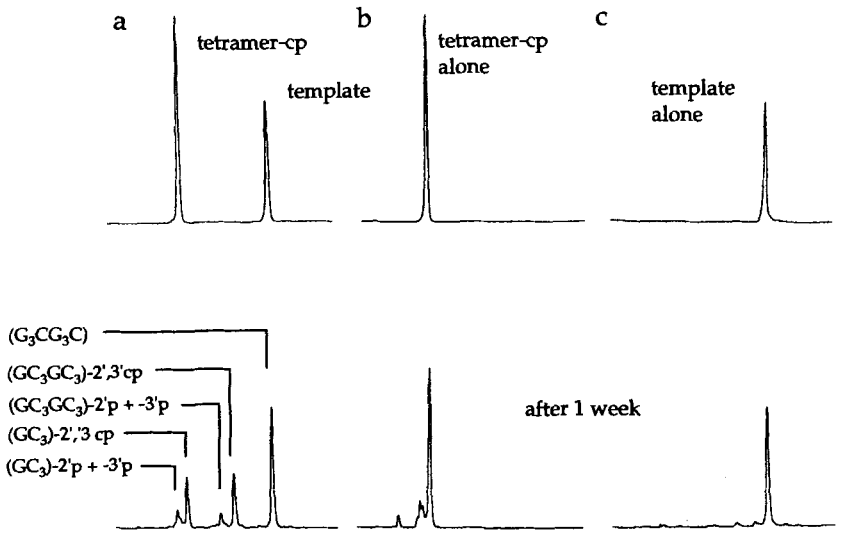


Fig. 5. *HPLC Traces after 1 week of* a) *ligation of 450* µM *pr(GCCC)-2',3'-cp in the presence of 150* µM *pr(GGGCGGGC) as template*; b) *450* µM *pr(GCCC)-2',3'-cp alone*; c) *150* µM *pr(GGGCGGGC) alone; each in 1.5*M *LiCl, 1 m*M *EDTA, 100 m*M *HEPES, pH 8.5, H_2O, room temperature.* Traces at the top refer to the start of the reaction.

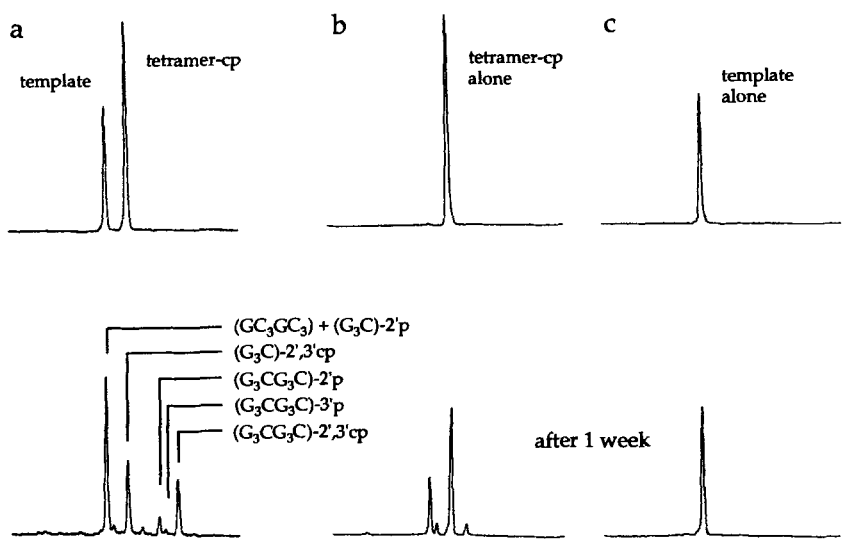


Fig. 6. *HPLC Traces after 1 week of* a) *ligation of 450* µM *pr(GGGC)-2',3'-cp in the presence of 150* µM *pr(GCCCGCC) as template*, b) *450* µM *pr(GGGC)-2',3'-cp alone*, c) *150* µM *pr(GCCCGCCC) alone; each in 1.5*M *LiCl, 1 m*M *EDTA, 100 m*M *HEPES, pH 8.5, H_2O, room temperature.* Traces at the top refer to the start of the reaction.

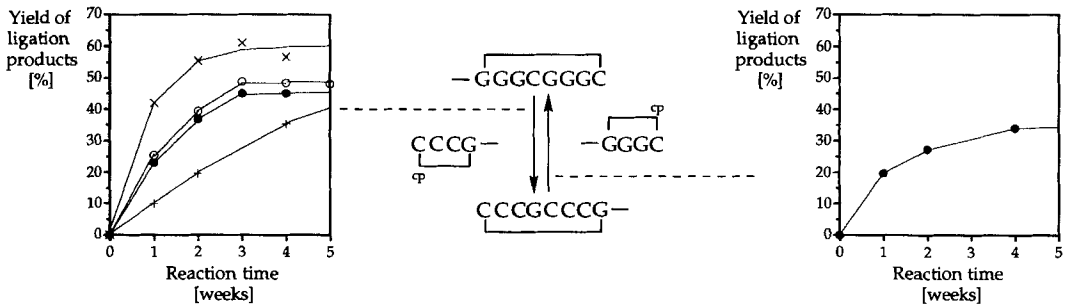


Fig. 7. *Plots of ligation product yields* vs. *reaction time of the ligation of 450* μM *pr(GCCC)-2',3'-cp in the presence of 150* μM *pr(GGGCGGGC) as template* (left) *and of 450* μM *pr(GGGC)-2',3'-cp in the presence of 150* μM *pr(GCCCGCCC) as template* (right) *in* (●): *1*M *NaCl, 25 mM* $MgCl_2$, *100 mM HEPES, pH 8.0,* H_2O; (+): *1*M *NaCl, 5 mM EDTA, 100 mM HEPES, pH 8.0,* H_2O; (○): *1*M *LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.0,* H_2O; (×): *1*M *LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5,* H_2O. All reactions at room temperature.

Table 1. *Reaction Conditions, Yields, and Ligation/Cyclophosphate Hydrolysis Rate Ratios of Ligation of 450* μM *pr(GCCC)-2',3'-p in the Presence of 150* μM *pr(GGGCGGGC) in 100 mM HEPES buffer* (pH as indicated, room temperature)

Entry	Electrolyte	Divalent cation	pH	Octamer yield [%][a])	Reaction time [weeks][b])	Rate ratio: ligation-to-cyclophosphate hydrolysis
1	1M NaCl	25 mM $MgCl_2$	8.0	45	4	~ 1:2
2	1M NaCl	25 mM $CaCl_2$	8.0	45	4	~ 1:2
3	1M NaCl	25 mM $MnCl_2$	8.0	22	1	< 1:5
4	1M NaCl	2.5 mM $MnCl_2$	8.0	33	4	~ 1:3
5	1M NaCl	–	8.0	48	10	~ 3:2
6	1M KCl	–	8.0	41 *	7 *	~ 3:2
7	1M Me_4NCl	–	8.0	5 **	11 **	~ 1:1
8	1M LiCl	–	7.0	27 ***	8 ***	~ 1:2
9	1M LiCl	–	8.0	48	4	~ 3:2
10	1M LiCl	–	8.5	60	3	~ 2:1
11	1M LiCl	–	9.0	60	1	~ 3:1
12	2M LiCl	–	8.0	57	1	~ 3:1
13	1.5M LiCl	–	8.5	57	1	~ 3:1

[a]) Reactions were monitored by HPLC, and yields were calculated with respect to the template. [b]) The reaction time is defined as the time after which no further increase of ligation products was observable. (*): Reaction no longer monitored after 7 weeks (HPLC trace comparable to *Entry 9* after 2 weeks; yield 42%); (**): after 11 weeks; (***): after 8 weeks.

seem to undergo the recyclization less reliably; however, definite information on their behavior is lacking.

The role of LiCl as a substitute for the combination of NaCl as bulk electrolyte and of $MgCl_2$ as catalyst in template-controlled ligations of tetramer-2',3'-cyclophosphates is of interest not only from a practical, but also from a mechanistic point of view. In

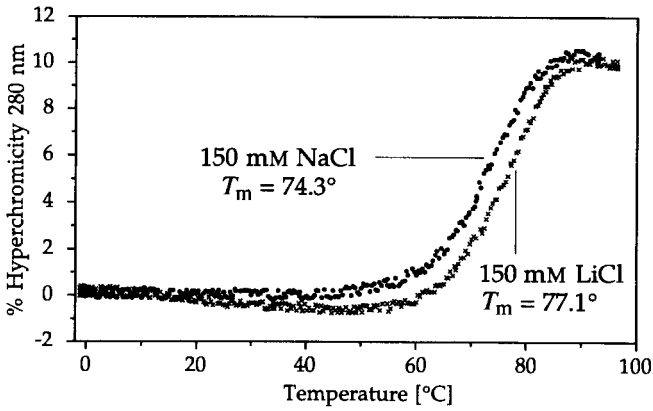


Fig. 8. *UV Melting curves of pr(GGGCGGGC) + pr(GCCCGCCC) 1:1.* (●): In 150 mM NaCl (each 3.4 μM); (×): in 150 mM LiCl (each 5.0 μM); both in 10 mM *Tris* · HCl, pH 7.0.

replacing the Mg^{2+} ion as a catalyst in both hydrolysis and alcoholysis (ligation) of the cyclophosphate group, Li^+ ion shifts the ratio of the two reactions in favor of ligation, a phenomenon which is probably related to differences in the structure and properties of the hydration shells of the two metal ions[9]).

Fig. 8 illustrates that T_m values of p-RNA duplexes in LiCl solutions tend to be somewhat – but certainly not much – higher than they are in NaCl solution under otherwise identical conditions.

Under all conditions tested, no ligation whatsoever was observed in the absence of template. Our tentative conclusion must be that the rate acceleration exerted by the template is enormous, but we have not quantified this statement. Not surprisingly, turnover of the template 'catalyst' in this standard series could not be enforced by simply applying a larger excess of ligand cyclophosphate, the thermal stability of the product duplex at the applied reaction temperature being too high.

4. Product Characterization and Regioselectivity of Ligation. – Detailed characterization and identification of ligation products were accomplished for the products of the transcription of pr(GGGCGGGC) into pr(GCCCGCCC)-2′,3′-cp and of the latter sequence (as the 2′-OH derivative) back into the 2′,3′-cyclophosphate of the former (*Figs. 3–7*). The experiments, in which the analytical data were collected, were run under the (originally developed) conditions (1M NaCl, 25 mM $MgCl_2$, 0.1M HEPES buffer at pH 8.0), in which cylophosphate hydrolysis is faster than ligation [4]. Products were separated by HPLC at pH 11.5 (where they run as single strands) and isolated after desalting by filtration through a *Sep-Pak*® cartridge as (mainly) triethylammonium salts.

---

[9]) The remarkable role Li^+ ion seems to play in this reaction is reminiscent of the beneficial role that Li^+ ion was found to play in an earlier investigation from our laboratory [20] (catalysis of sulfide contractions of thioesters by $LiClO_4$ or LiBr in organic solvents). See the reviews of *Loupy* and *Tchoubar* [21], and *Seebach et al.* [22].

Product fractions containing the hydrolyzed phosphates at the 2′-terminus were isolated as mixtures of mostly 2′-phosphates containing (varying) minor amounts of 3′-phosphates. Constitutional assignments of the products of the two ligations are based on the following data (see also [4]).

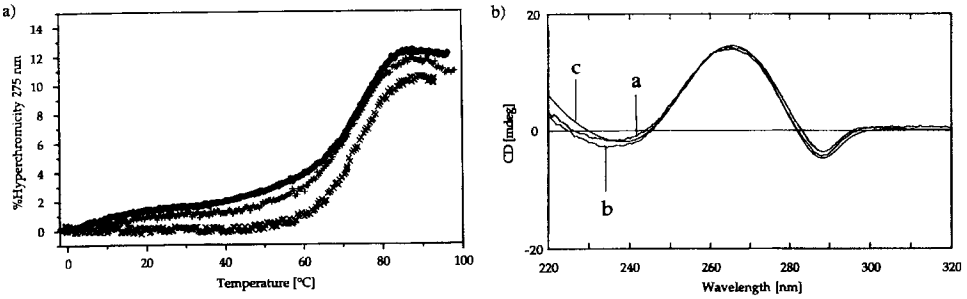


Fig. 9. *Identification of ligation products by UV and CD spectroscopy. a)* UV Melting curves of (×): pr(GCCCGCCC) + pr(GGGCGGGC) 1:1 (authentic sequences), 2.7 μM each, T_m = 75.3°; (●) pr(GCCCGCCC)-2′/3′-p (*ex* ligation) + pr(GGGCGGGC) (auth.) 1:1, 2.3 μM each, T_m = 75.1°; (+): pr(GCCCGCC) (auth.) + pr(GGGCGGGC)-2′-p (*ex* ligation) 1:1, 1.0 μM each, T_m = 74.2°; in 150 mM NaCl, 10 mM *Tris* · HCl, pH 7.0; b) CD spectra: curve *a*: pr(GCCCGCCC)-2′-p/3′-p (*ex* ligation) + pr(GGGCGGGC) (auth.) 1:1, 2.3 μM each; curve *b*: pr(GCCCGCCC)-2′-p (auth.) + pr(GGGCGGGC) (auth.) 1:1, 2.3 μM each; curve *c*: pr(GCCCGCCC) (auth.) + pr(GGGCGGGC)-2′-p (*ex* ligation) 1:1, 1.9 μM each; in 150 mM NaCl, 10 mM *Tris* · HCl, pH 7.0, 25°.

Table 2. *Calculated and Observed Mass for Ligation Products from Ligation of pr(GC_3)-2′,3′-cp in the Presence of pr(G_3CG_3C), and of pr(G_3C)-2′,3′-cp in the Presence of pr(GC_3GC_3).* Mass determined by MALDI-TOF mass spectrometry (*cf. Pieles et al.* [23]).

Ligation product	$[M - H]_{calc.}$	$[M - H]_{obs.}$
pr(GC_3GC_3)-2′- and -3′-p	2538.5	2536.2
pr(GC_3GC_3)-2′,3′-cp	2520.5	2519.9
pr(G_3CG_3C)-2′-p	2698.6	2697.4
pr(G_3CG_3C)-2′,3′-cp	2680.6	2680.8

MALDI-TOF Mass spectrometry [23] of product fractions (*Table 2*)[10]) confirmed sequence length and base content, as well as hydrolysis status. UV Melting curves of 1:1 mixtures of product fractions and their respective templates (*Fig. 9, a*), as well as CD spectra of these mixtures (*Fig. 9, b*), in comparison with the corresponding data for the authentic template-template duplex, confirm base-sequence complementarity. Treatment of the major product fractions (mostly 2′-phosphate) with alkaline phosphatase converted them to corresponding dephosphorylated fractions which were identified by HPLC co-injection with the authentic octamers pr(GCCCGCCC) and

[10]) We thank Dr. *U. Pieles*, *Ciba-Geigy AG*, Basel, for determining all MALDI-TOF mass spectra in his laboratory.

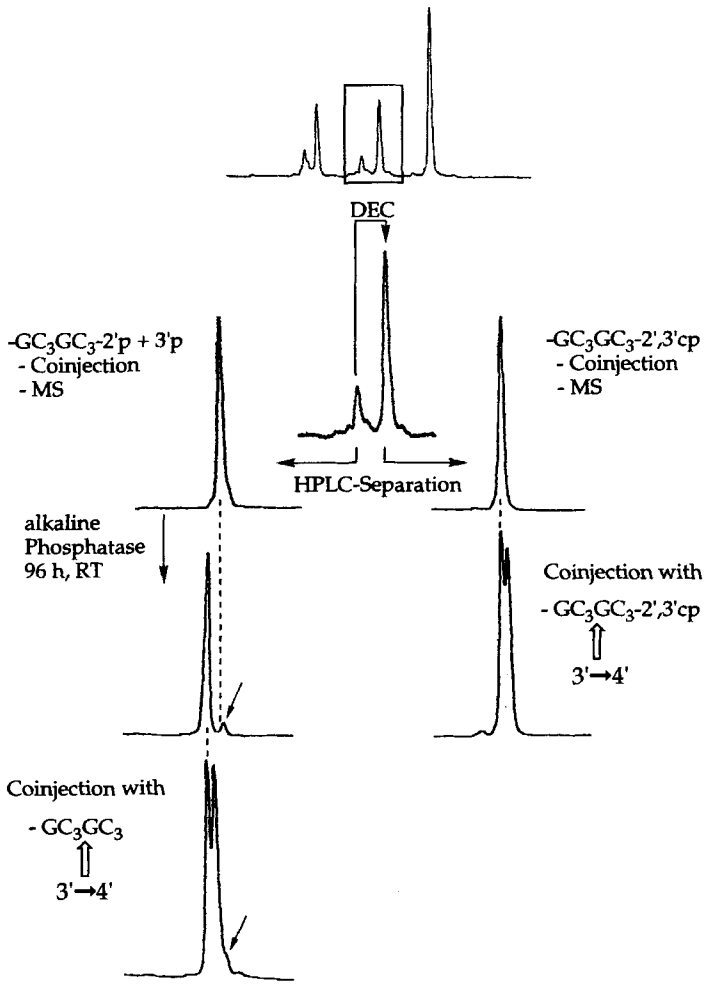


Fig. 10. *Identification of products of the ligation of pr(GCCC)-2',3'-cp in the presence of pr(GGGCGGGC).* Top HPLC trace: reaction in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, H_2O, room temperature after 1 week. Reaction conditions for interconversions: pr(GC_3GC_3)-2'-p + -3'-p → pr(GC_3GC_3)-2',3'-cp: excess DEC in H_2O, room temperature, 2 h; pr(GC_3GC_3)-2'-p + -3'-p → pr(GC_3GC_3): 2500 U/ml alkaline phosphatase (from calf intestine, EC 3.1.3.1), in 10 mM diethanolamine · HCl, pH 9.8, room temperature, 96 h. For coinjections, see text.

pr(GGGCGGGC), respectively (*Fig. 10*)[11]). Finally, treatment of the ligation product phosphate fractions with DEC (excess, in H_2O, room temperature, 2 h) converted them

[11]) The 3'-phosphate of pr(GCCCGCCC) has been observed to be dephosphorylated by alkaline phosphatase much more slowly than the corresponding 2'-phosphate: it remained almost unchanged during the reaction time needed for complete dephosphorylation of the 2'-phosphate (axial *vs.* equatorial OH). The 3'-phosphate was identified by co-injection of a partially dephosphorylated mixture with authentic 3'-phosphate. However, a 1:1 mixture of authentic pr(GCCCGCCC)-2'-p and pr(GCCCGCCC)-3'-p appeared in HPLC as a broadened peak which could not be separated.

to 2′,3′-cyclophosphate fractions which were identified in HPLC by co-injection with the product fractions pr(GCCCGCCC)-2′,3′-cp and pr(GGGCGGGC)-2′,3′-cp, respectively.

The question as to what extent the ligation process is regioselective refers to a side reaction that would produce (4′ → 3′) phosphodiester junctions at the ligation site in addition to the correct (4′ → 2′) junction (*cf. Fig. 11* and discussion below). We have checked this possibility for the transcription of pr(GGGCGGGC) by synthesizing the potential side products pr(GCCC$^{3',4'}$GCCC)-2′,3′-cp and -2′-p (as well as the dephosphorylated derivative) containing the critical (4′ → 3′) junction, and determining their HPLC behavior under the conditions of the ligation product analysis[12]). The HPLC trace of the co-injection of the dephosphorylated ligation product with authentic pr(GCCC$^{3',4'}$GCCC) clearly shows two separated peaks as does the trace of the co-injection of the ligation product pr(GCCCGCCC)-2′,3′-cp with authentic pr(GCCC$^{3',4'}$GCCC)-2′,3′-cp (*Fig. 10*). There is no discernible amount of the (4′ → 3′) regioisomer of the sequence pr(GCCCGCCC) formed in the transcription of pr(GGGCGGGC)[13]). This demonstrates that it is exclusively the 3′-O-center of the ligand's cyclophosphate group that acts as the leaving group in the transesterification reaction.

Model considerations on the levels both of idealized conformations [18] [4] [24] and of molecular-dynamics modeling [1] of p-RNA duplexes have predicted template-directed ligation to proceed with unambiguous regioselectivity to give (4′ → 2′) phosphodi-

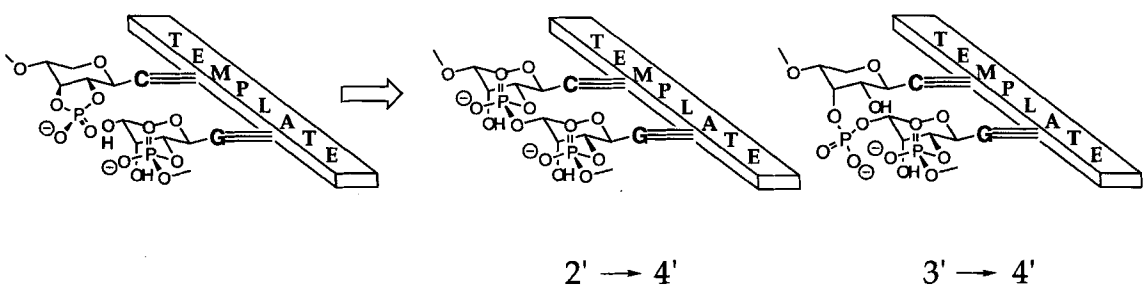


Fig. 11. *Template-controlled ligation of tetramer-2′,3′-cyclophosphates.* Does it lead regioselectively to (2′ → 4′) phosphodiester junctions or also to isomeric (3′ → 4′) junctions?

---

[12]) The synthesis of the central (4′ → 3′) phosphodiester junction involved as phosphoramidite building block the 4′-(dimethoxytrityl)-2′-benzoyl- (instead of 3′-benzoyl)-cytosine derivative (for details, see [4] and [17]). Surprisingly, upon mixing of equivalent amounts of the regioisomer pr(GCCC$^{3',4'}$GCCC) with pr(GGGC$^{2',4'}$GGGC), a duplex is formed (T_m = 74°, 3.4 μM) that melts almost as high as the normal (4′ → 2′) duplex (T_m = 76°, 3.4 μM), but with distinctly lower UV hyperchromicity. In fact, the octamer sequence pr(GC$_3^{3',4'}$GC$_3$) was observed to induce ligation of pr(G$_3$C)-2′,3′-cp quite efficiently: 18.7% after 5 weeks (7.1% after 1 week) of pr(G$_3$CG$_3$C)-2′,3′-cp (p) (1.0M NaCl, 0.025M $MgCl_2$, pH 8.0, room temperature). Exposing the duplex pr(GC$_3^{3',4'}$GC$_3$) · pr(G$_3$CG$_3$C) to the above conditions did not lead to intrastrand rearrangement of pr(GC$_3^{3',4'}$GC$_3$) into pr(GC$_3^{2',4'}$GC$_3$) during 9 weeks, according to HPLC analysis under denaturing conditions.

[13]) The small HPLC signal on the right of the main signal in the post-phosphatase treatment product (to the left of *Fig. 11*) is not the (4′ → 3′) regioisomer (according to t_R), but (very probably) residual 3′-phosphate.

ester junctions (*Figs. 11* and *12*)[14]). This expectation is based on p-RNA's most characteristic structural feature, namely, the large inclination between (averaged) backbone and base-pair axes [1] [18]. In the preligation complex, the two ligation partners position themselves on the template strand in such a constellation[15]) that the free OH group at the 4′-end of one partner and the cyclophosphate substitution center of the other partner are colinear with the 3′-O-center of the cyclophosphate-bearing partner[16]). With such a preorganization of reaction centers, the transesterification reaction is expected to take the path in which the 3′- (and not the 2′)-O-function of the cyclophosphate-bearing partner acts as the leaving group. This corresponds to a kinetic selection of the (4′ → 2′) phosphodiester junction over the isomeric (4′ → 3′) junction.

The topological and mechanistic solidity of this model-based argumentation, as well as its concurrence with the regioselectivity documented experimentally for the case described above, justifies – so we think – the extrapolation and (tentative) conclusion according to which regioselectivity in the formation of phosphodiester (4′ → 2′) junctions should be a sequence-independent property of template-controlled ligation in the p-RNA series.

5. Chiroselectivity of Ligation. – Since ligation of p-RNA tetramer-2′,3′-cyclophosphates under all conditions tested so far seems to depend strictly on the presence of a

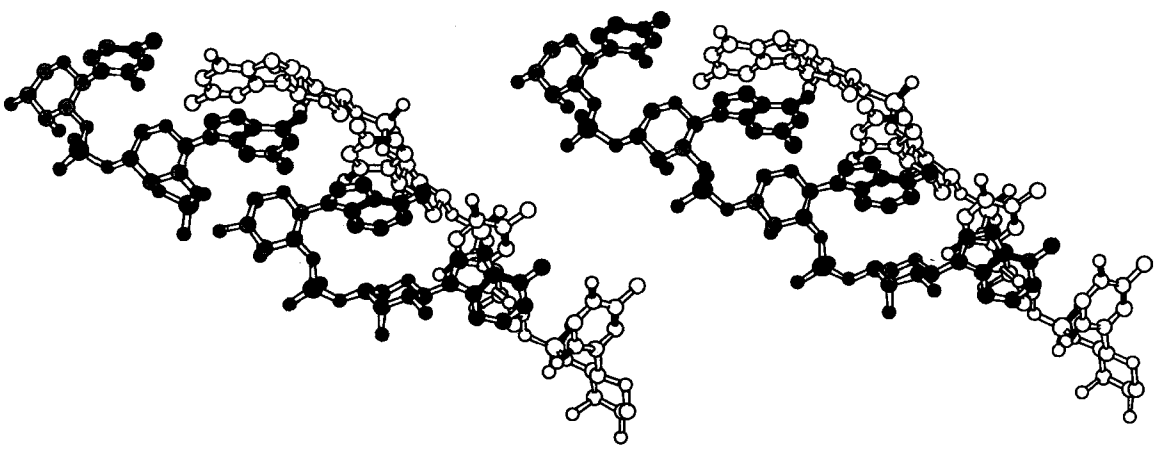


Fig. 12. *Modeling of the pre-ligation complex of a template-controlled ligation in the p-RNA series* (courtesy of Dr. *R. Wolf*, *Ciba-Geigy AG*, Basel). For the structure of a p-RNA duplex, see [1].

---

[14]) *Fig. 12* presents a modelling result originating in the work of *R. M. Wolf*, *Ciba-Geigy AG*, Basel (see [1]). We thank Dr. *Wolf* for allowing us to publish the figure in this paper.

[15]) We use the term 'constellation' in the following sense: relative positioning in space of non-covalently interacting molecules in supramolecular complexes (see comments in [25]).

[16]) A similar co-linearity argument (see also [26]) has been used by *Usher* [12] in interpreting the formation of the (2′ → 5′) phosphodiester junction in his template-controlled ligation to an RNA oligomer (see *Chapt. 10*).

template and does not occur to any discernible extent without one, it can be expected that it also is chiroselective[17]). In their work on template-directed oligomerization of guanosine-5′-phospho-2-methylimidazole, *Joyce et al.* [27] showed that oligomerization of the activated L-guanosine derivative on a D-poly(C)-template is 'far less efficient' than that of the D-guanosine derivative, and, furthermore, that the presence of the L-enantiomer in the reaction mixture substantially reduces the efficiency of the oligomerization of the D-enantiomer. The reason for the latter phenomenon ('enantiomeric cross inhibition' [27]) was proposed by the authors to reside in the ability of the enantio-nucleotide

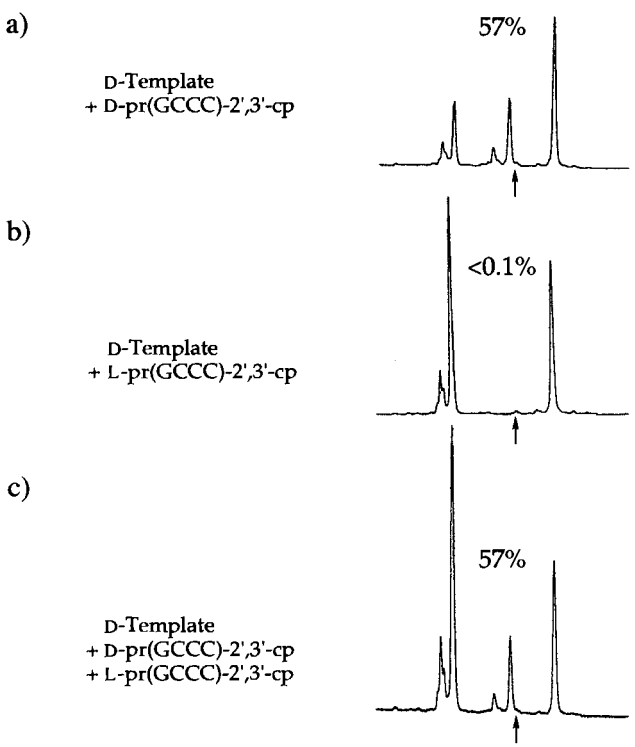


Fig. 13. *Chiroselectivity of ligation.* HPLC Traces after 1 week of the ligation of *a)* 450 μM D-pr(GCCC)-2′,3′-cp in the presence of 150 μM D-pr(GGGCGGGC); *b)* 450 μM L-pr(GCCC)-2′,3′-cp in the presence of 150 μM D-pr(GGGCGGGC); *c)* 450 μM D-pr(GCCC)-2′,3′-cp and 450 μM L-pr(GCCC)-2′,3′-cp in the presence of 150 μM D-pr(GGGCGGGC); all reactions in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, H_2O, room temperature. Arrow (↑) does not refer to a ligation product, but to an unidentified template decomposition product. Yields after 2 weeks were: *a)* 57%, *b)* < 0.1%, *c)* 60%.

[17]) The stereoselectivity we are dealing with is enantioselectivity in some of the cases (referring to sense of chirality of a homochiral template and a homochiral ligand) and diastereoselectivity in others (referring to the sense of chirality of ribopyranosyl units within a heterochiral template or heterochiral ligand). To specify and differentiate the latter cases from conventional diastereoselectivity (referring to the configuration at *any* stereocenter in backbone units), we use the term '*chiroselectivity*', implying that the selectivity refers only to the diversity in sense of chirality of backbone units within an oligonucleotide sequence.

monomer to undergo base-pairing with the template by making use of a *syn*-conformation between nucleobase and ribofuranosyl ring. When the chiroselectivity of the monomer-oligomerization process is less than perfect, then incorporation of L-units into the growing oligomer chain can occur, and this can lead to chain termination [27]. As far as we know, the degree of chiroselection that can be achieved in template-controlled ligation of small oligomers in oligonucleotide systems, and the optimal sequence length of ligand oligomers for maximal chiroselection, have not been studied until now.

No ligation whatsoever was observed, when we subjected the L-enantiomer of pr(GCCC)-2′,3′-cp (all ribopyranosyl units being L) to reaction conditions under which otherwise the template D-pr (GGGCGGGC) is transcribed in 57% yield to the complementary sequence D-pr(GCCCGCCC)-2′,3′-cp (p's) (*Fig. 13, a* and *b*). Furthermore, no inhibition of the transcription of the D-template into the complementary D-sequence was found when the reaction was run in the presence of equimolar amounts of D- and L-ligands pr(GCCC)-2′,3′-cp (*Figs. 13, c*, and *14*). This agrees with previous findings in the p-RNA series [3], according to which heterochiral pairing between complementary sequences of enantiomeric oligomers containing canonical bases is either absent or, compared to homochiral pairing, very weak.

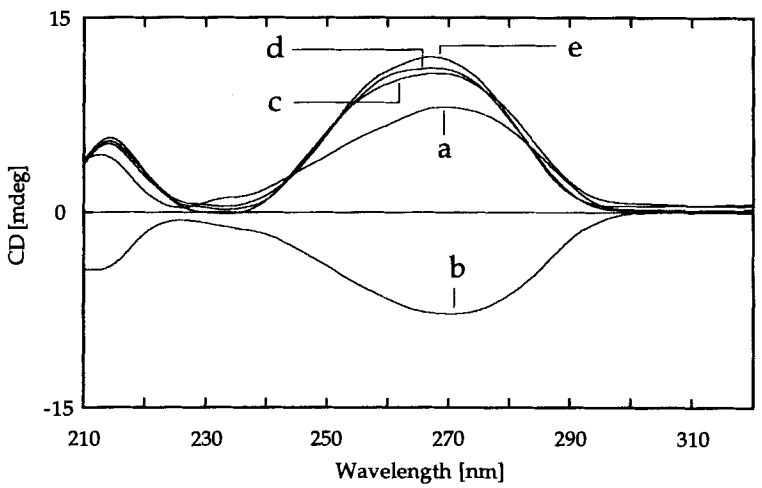


Fig. 14. *CD Spectra.* Curve *a*: D-pr(GCCC)-2′,3′-cp (authentic); curve *b*: L-pr(GCCC)-2′,3′-cp (auth.); curve *c*: D-pr(GCCCGCCC)-3′-p (auth.); curve *d*: pr(GCCCGCCC)-2′,3′-cp (*ex* ligation, *Fig. 13, c*); curve *e*: pr(GCCCGCCC)-2′-p (*ex* ligation, *Fig. 13, c*, contains minor amounts of the corresponding -3′-p). 11 μM oligonucleotide in 150 mM NaCl, 10 mM *Tris* · HCl, pH 7.0, room temperature.

Fig. 15 presents the observations made in a series of four experiments in each of which a heterochiral diastereoisomer of D-pr(GCCC)-2′,3′-cp was subjected to standard ligation conditions (see *Chapt. 3*), each time with just one of the four ribopyranosyl units as L-unit. Whereas in three of these experiments (*Fig. 15, b, c*, and *d*) no ligation was detectable within 1 week (less than 0.1% yield), in the experiment with the guanine-

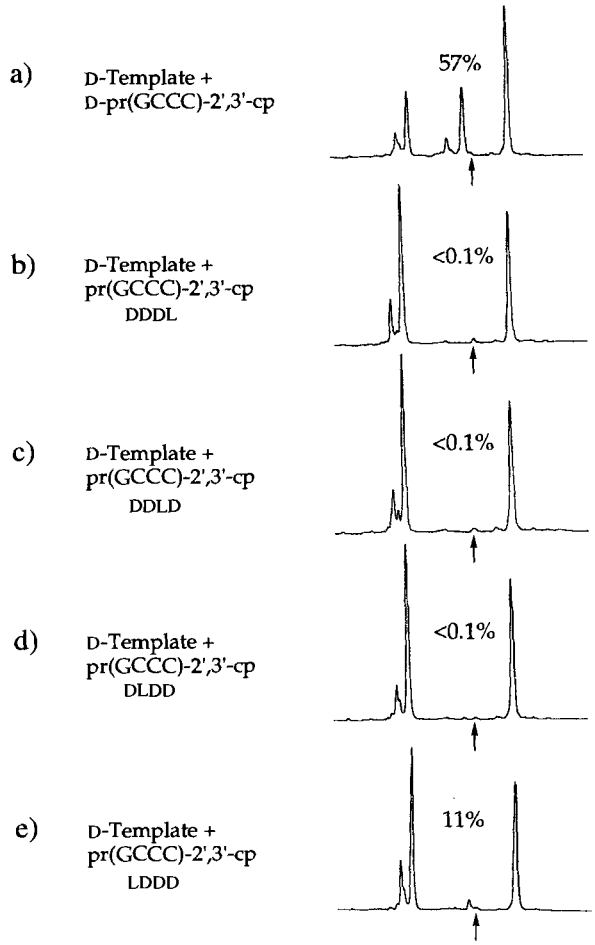


Fig. 15. *Chiroselectivity of ligation.* HPLC Traces after 1 week of five ligation experiments (*a–e*) with D-pr(GGGCGGGC) as template in the presence of heterochiral diastereoisomers of D-pr(GCCC)-2',3'-cp, in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, room temperature; configuration of chiro-diastereoisomers as indicated below the sequence formula. Arrow (↑) does not refer to a ligation product, but to an unidentified template decomposition product. Yields after 2 weeks: *a*) 57%, *b*) < 0.1%, *c*) < 0.1%, *d*) < 0.1%, *e*) 14%.

bearing unit at the 4'-end replaced by the L-unit, the ligation rate (as judged from the yield of octamer fractions after 1 week) amounted to *ca.* 10% of the rate of ligation with the homochiral ligand (*Fig. 15, e*). The ligation turns out to be highly chiroselective, except when the L-unit is located at the 4'-end of the tetramer-2',3'-cyclophosphate and carries a purine as nucleobase. In [6], we described our more recent work on the chiroselectivity of self-directed oligomerization of hemi-self-complementary tetramer-2',3'-cyclophosphates to higher p-RNA oligomers, in which an analogous chiroselectivity pattern is observed, and where we have found that the remarkable drop in chiroselectivity for (D → L) substitutions at the 4'-end is associated with purines only and not pyrimidines.

The reader is referred to [6] for a discussion and qualitative rationalization of the observed chiroselection phenomena.

6. Fidelity and Sequence Dependence of Ligation. – Transcription of nucleic-acid base sequences by template-controlled ligation of oligonucleotides (as opposed to oligomerization of mononucleotides) is expected to be confronted with problems of transcription fidelity [8]. The frequency and severity of such problems will depend, among other factors, on the length of the ligand sequence (the shorter the sequence the higher the fidelity), on the related mismatch sensitivity of pre-ligation complexes (the higher the better), and on the gap between the rates of spontaneous *vs.* template-induced ligation (the broader the better), with the consequent implication for the degree of sensitivity of the ligation transition state toward template-induced pre-organization of the ligand's reaction centers (the higher the better). Mildness of ligand phosphate activation must be a crucial factor in this context; from this point of view, the type of activation adapted for p-RNA tetramers may well be a fortunate choice, since no ligation whatsoever has ever been observed to occur in the absence of template and under conditions, where the template-controlled reaction has, apart from cyclophosphate-group hydrolysis, essentially no competition.

The greater part of the experiments run to test transcription fidelity of p-RNA sequences employed the original ligation conditions (1M NaCl, containing 25 mM $MgCl_2$ at pH 8.0) under which cyclophosphate-group hydrolysis competes more efficiently than under the conditions adopted later (see *Table 1*); however, there are no obvious reasons for assuming that the conclusion reached from experiments under the former conditions would not also be valid for the latter.

Fig. 16 summarizes the ligation yields observed (after 4 weeks) in experiments (*A*–*T*) where six different G- and C-containing octamer sequences (with free 4′- and 2′-ends) were left to induce ligation of a set of fully complementary (*Exper. A*–*D*, *N*, and *R*), or only partially complementary (*Exper. E*–*H*, *I*–*M*, *O*–*Q*, *S*, and *T*) tetramer-2′,3′-cyclophosphates under $MgCl_2$ conditions (see caption of *Fig. 16*). The mismatch-free ligations *A*–*D* reveal a sequence dependence of ligation yields. These yields are thought to result from the competition between the rate of cyclophosphate-group hydrolysis (considered to be only weakly dependent on base sequence) and the rates of ligation. Since the base sequences in the product duplexes of *Exper. A* and *B* are identical, as they are for *C* and *D*, and since both product duplexes are held together by the same number of *Watson-Crick* G·C pairs, the differences in ligation yields in these four experiments are interpreted to reflect the differences in interstrand base stacking at the ligation sites (see discussion in *Chapt. 9*).

In *Exper. E*–*H*, the four tetramer ligands used in *Exper. A*–*D* are interchanged and combined with the four template sequences in such a way that in each case the ternary pre-ligation complex is disturbed by four mismatches (see arrows). Not unexpectedly, in none of the four experiments was there any ligation discernable. Since the mismatches are located in the central region of the tetramer rather than at the ligation sites, the reason for the absence of ligation is to be seen in a destabilization of pre-ligation complexes.

Still another possibilty for combining the four tetramer ligands with the four templates is the theme of *Exper. I*–*M*. The ternary pre-ligation complexes, rather than showing mismatches, are composed of seven (and not eight) *Watson-Crick* G·C pairs

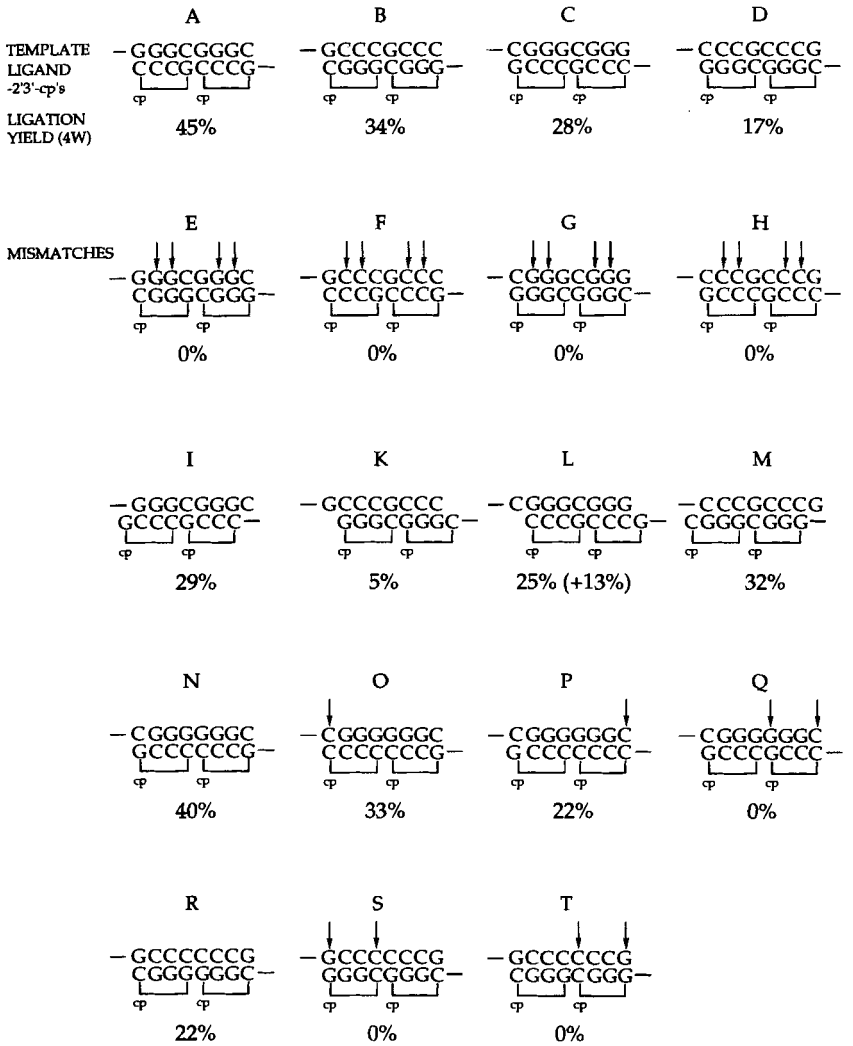


Fig. 16. *Sequence dependence and fidelity of ligative replication.* Reaction conditions: 1M NaCl in the presence of 0.025M $MgCl_2$, pH 8.0, 1 mM HEPES, room temperature, *c* of template 150 μM, *c* of tetramers 2 × 225 μM (= conditions used originally; unfavorable ligation-to-hydrolysis ratio!). Yields: complementary octamers (2′,3′-cyclophosphates + 2′- and 3′-phosphates) formed after 4 weeks as determined by HPLC (see *Fig. 4*). *Exper. L*: yield in brackets refers to two additional HPLC peaks indicating the formation of a dodecamer and (presumably) a hexadecamer. Vertical arrows indicate mismatches in pre-ligation complexes. For HPLC documentation of experiments *C*, *D*, *G*, *H*, *L*, *N*, and *R*, see *Exper. Part.*

with single dangling bases at either the 2′- (*I* and *M*) or the 4′-end (*K* and *L*) of the ligand strands. With the exception of *Exper. K*, the ligations in this series are remarkably efficient. In *Exper. L*, the reaction leads, in addition to the octamer, to fractions which are the dodecamer (9%) and (presumably) the hexadecamer (4%). As will be discussed in more detail in *Chapter. 9*, this special behavior again is interpretable in terms of

differences in interstrand stacking at ligation sites. The fact that the four template sequences can be transcribed in two different modes is not necessarily to be seen as an example of lack of fidelity of transcription; it is a consequence of the sequence isomerism and repetitivity in this special templates.

The pair of *Exper. N* and *R* represent a replicative transcription of an octamer sequence which requires two different tetramer ligands, rather than a single one twice. In *Exper. O–Q*, *S*, and *T*, the same two templates were combined with ligand sequences in such a way that there are one or two mismatches at either end, or at ligation sites. *Exper. O* and *P* demonstrate the effect of a single mismatch at either the 4′- or 2′-end of the template (three base positions away from the ligation site): it hampers ligation only to a marginal extent. If, however, such mismatches occur (within the same ligand tetramer) concomitantly with a mismatch at the ligation site, ligation is strongly impaired.

The results obtained present a few elements of a mismatch analysis which remains far from being complete. Nevertheless, together with the observations on the transcription experiments involving the standard sequence pr(GGGCGGGC) and the chiro-mismatches presented in *Fig. 15*, they point to the degree of the fidelity of transcription with respect to the ligation sites, as well as to the sequence positions in the close neighborhood thereof[18]).

7. Ligative Replication of the Dodecamer Sequence pr(GGGCGGGCGGGC). – So far, we have described examples of sequence copying in which octamer sequences mediate the ligation of two tetramers to a complementary octamer sequence. We now summarize the results obtained in ligation experiments with the two complementary template sequences pr(GGGCGGGCGGGC) and pr(GCCCGCCCGCCC) using corresponding tetramer-2′,3′-cyclophosphates as ligands (*Figs. 17–19*). Furthermore, we report observations made in exploratory attempts to use 2′,3′-cyclophosphates of dimers for copying the same two dodecamer sequences by elongation of a corresponding primer sequence (*Figs. 20–22*).

Neither of the two dodecamer template sequences exists in solution at ambient temperature as a single strand; this is evident from the UV melting curves of the two individual sequences shown in *Fig. 17,a*. The constitution of the self-pairing complexes has not been investigated; we conjecture that they are duplexes of the type shown, held together with six (G·C) *Watson-Crick* pairs twice interrupted by two consecutive (G·G) mismatches in one case, and by two (C·C) mismatches in the other[19]).

---

[18]) How the notorious RNA mismatch guanine-uracil (thymine) operates in the p-RNA series has recently been tested under standard conditions (1.5M LiCl, pH 8.5): the template pr(GGGCGGGC) ligates the mismatch ligand pr(GTCC)-2′,3′-cp about half as fast than it ligates pr(GCCC)-2′,3′-cp (26 *vs.* 58% yield within a week). In the more sensitive test of self-directed oligomerization of p-RNA tetramer-2′,3′-cyclophosphates oligomerization [6], rate retardation in pr(GTCG)-2′,3′-cp and pr(GCTG)-2′,3′-cp in comparison to pr(GCCG)-2′,3′-cp is very much higher.

[19]) Corresponding hairpin structures (see [5]) containing a loop consisting of six bases and held together by two (G·C) base pairs are not excluded (dependence of T_m values on concentration not tested). On the other hand, a guanine tetrade for the case of the G-rich dodecamer is improbable (see [4]).

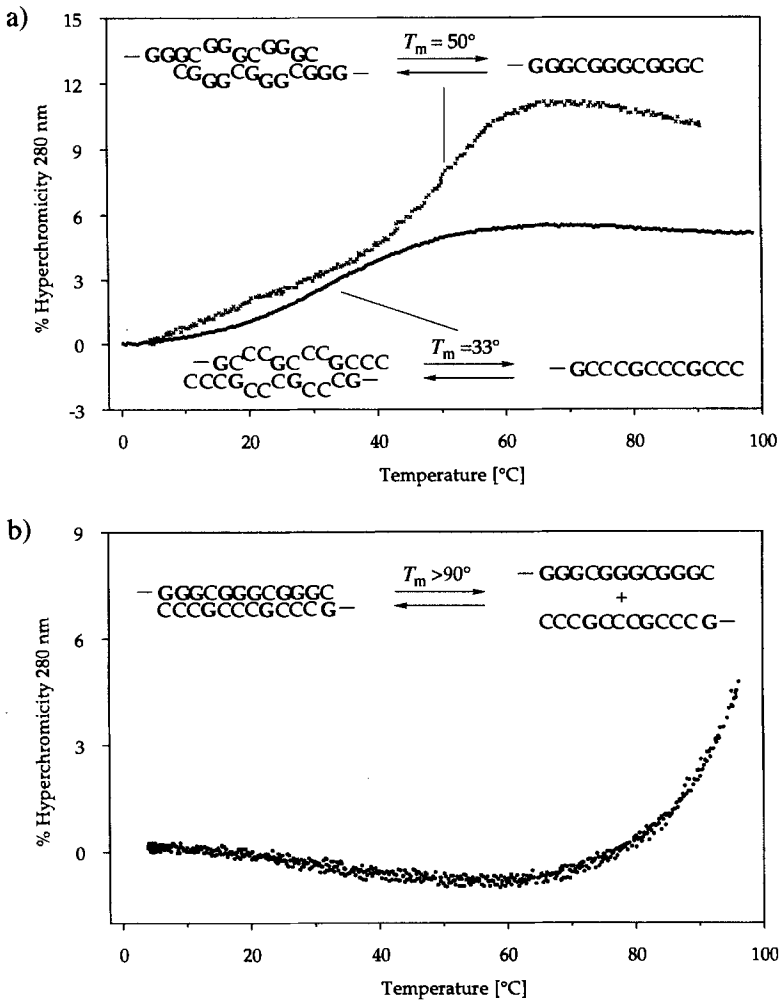


Fig. 17. *UV Melting curves of* a) *pr(GGGCGGGCGGGC)* (5.0 μM) *and pr(GCCCGCCCGCCC)* (5.2 μM); b) *pr(GGGCGGGCGGGC) + pr(GCCCGCCCGCCC) 1:1* (2.5 μM each). In 150 mM NaCl, 10 mM *Tris* · HCl, pH 7.0.

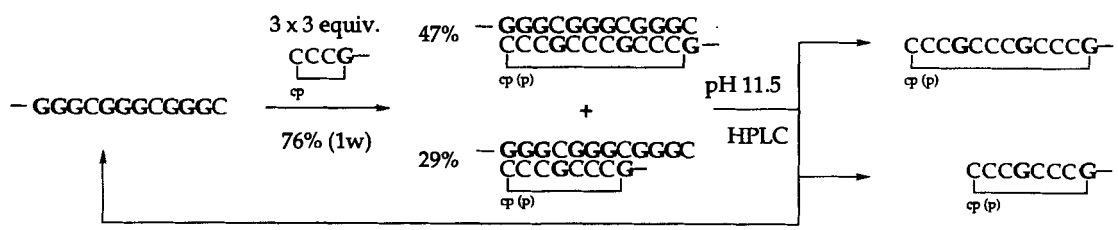
Mixing the two dodecamers in a 1:1 ratio leads to what (presumably) is a 1:1 duplex with a T_m value higher than 90° ($c = 2.5 + 2.5$ μM; *Fig. 17, b*)[20]).

Under standard conditions, the G-rich dodecamer is transcribed to the complementary C-rich dodecamer in 56% yield within 2 weeks (accompanied by 28% of the corresponding octamer) when a threefold excess of the tetramer pr(GCCC)-2′,3′-cp is

[20]) The origin of the consistently observed weak structuring of the melting curve at lower temperature has not been investigated. It may reflect an intermolecular parallel aggregation of duplex double strands by orthogonal pairing modes of a type we have tentatively formulated in [3].

used as a ligand (*Fig. 18*). The result corroborates what one should expect with regard to the ability of tetramer ligands to break up template self-pairing complexes as long as the latter have a mismatch handicap. Transcription of the C-rich transcript sequence back to the G-rich template sequence proceeds comparably well, yet expectedly in slightly lower yield under the same conditions (*Fig. 19*).

The constitutional assignments of the reaction products formed in these transcription experiments rest on their HPLC retention times (at pH 11.5) relative to those of the corresponding authentic dodecamers and octamers (see *Figs. 18, b*, and *19, b*).



b)

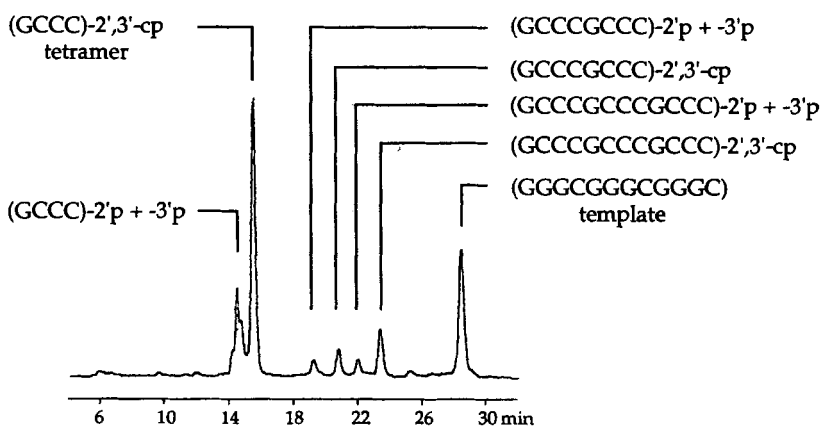


Fig. 18. *Reaction scheme* (a) *and HPLC trace after 1 week* (b) *of the ligation of 1.35 mM pr(GCCC)-2',3'-cp in the presence of 0.15 mM pr(GGGCGGGCGGGC) as template.* Reaction in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, room temperature. Yields after 2 weeks were: 28% pr(GCCCGCCC)-2'-p, -3'-p + -2',3'-cp, 56% pr(GCCCGCCCGCCC)-2'-p, -3'-p + -2',3'-cp.

Fig. 20 summarizes what we have observed in experiments addressing the question whether 2',3'-cyclophosphates of dimers could be used in place of tetramers as activated ligands in replicative ligations. The experiments consist in observing the formation of primer elongation when the dodecamer template pr(GGGCGGGCGGGC) is exposed to an excess of a 1:1 mixture of the dimers pr(CC)-2',3'-cp and pr(GC)-2',3'-cp in the presence of 1 equiv. (relative to template) of the octamer pr(GCCCGCCC) (free 2'-OH) as a primer at the template's 4'-end (for the constitutional assignment of the template-

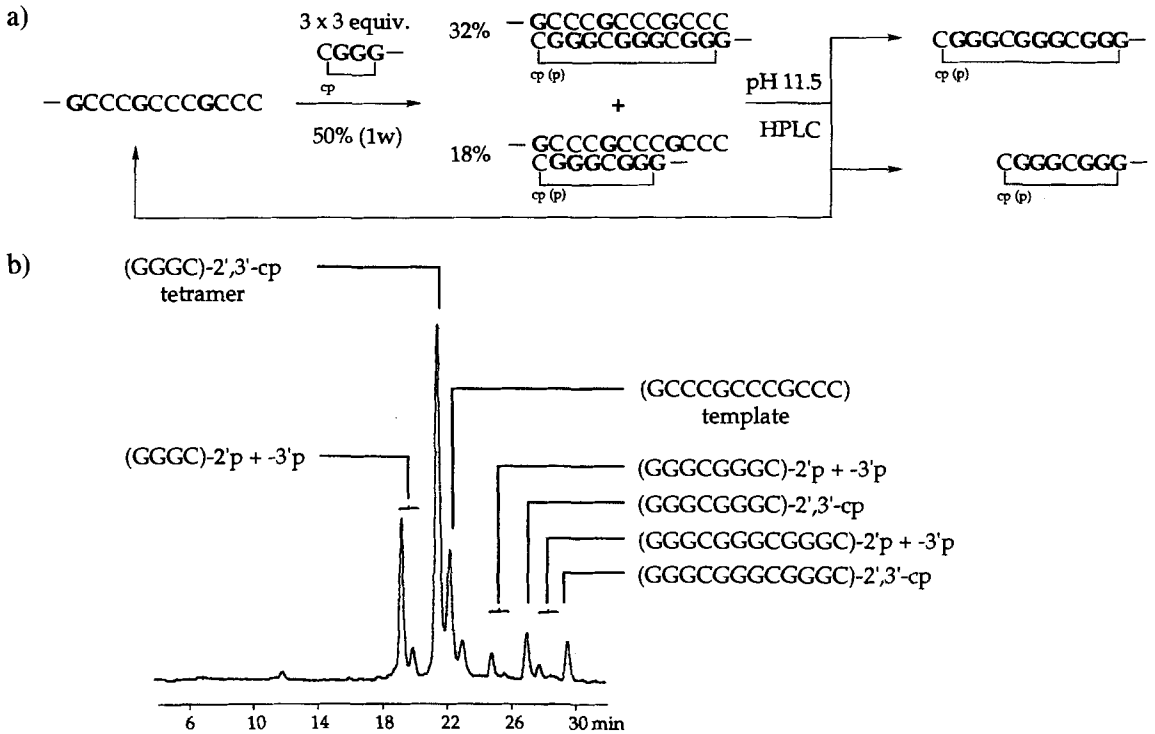


Fig. 19. *Reaction scheme* (a) *and HPLC trace after 1 week* (b) *of the ligation of 1.35 mM pr(GGGC)-2',3'-cp in the presence of 0.15 mM pr(GCCCGCCCGCCC) as template.* Reaction in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, room temperature. Yields after 2 weeks were: 36% pr(GGGCGGGC)-2'-p, -3'-p + -2',3'-cp, 32% pr(GGGCGGGCGGGC)-2'-p, -3'-p + -2',3'-cp.

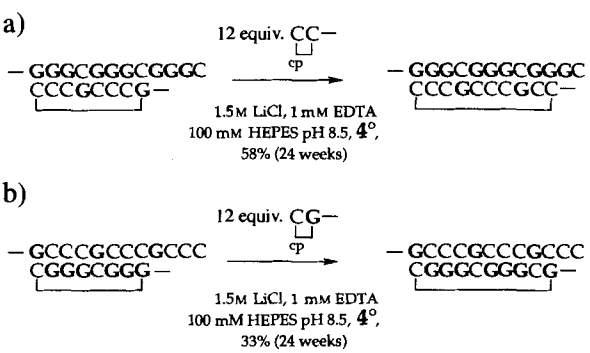


Fig. 20. *Elongation of octamer primers on dodecamer templates by ligation with dimer-2',3'-cyclophosphates.* For constitutional assignment of template primer duplexes, see *Fig. 29.*

primer duplex, see *Chapt. 9*). The expected lower stability of pre-ligation complexes in the case of dimer, as compared to tetramer ligands, as well as observations made in exploratory experiments, dictated the use of a reaction temperature below room temperature for

these elongations. The experiment carried out at 4° showed: first, elongation of the primer sequence does take place very slowly but also very cleanly; and, second, elongation stops at the decamer, the dimer pr(CC)-2′,3′-cp becoming incorporated, but not the pr(CG)-2′,3′-cp unit which would become the 2′-end of the dodecamer transcript. This latter result becomes plausible when its reason is analyzed (see *Chapt. 9*); therefore, the elongation experiment was rerun in the presence of the same primer and of pr(CC)-2′,3′-cp as the single dimer component. As before, the elongation for the decamers is slow but clean, as the HPLC traces over 24 weeks reaction time (*Fig. 21*) demonstrate. The rate of the corresponding reaction in the *retro*-series, where the dimer pr(GC)-2′,3′-cp is to be incorporated (*Fig. 22*), proceeds similarly, but still more slowly, as seems to be the general rule with relative rates when (C → G) *vs.* (G → C) transcriptions in the p-RNA series are compared.

The organizing role of the primer in these reactions seems quite essential: In an exploratory experiment using the G-rich dodecamer as template and a mixture of pr(CC)-cp and pr(GC)-cp (18 equiv., 2.7 mM, each) without primer, only a very slow formation of a product tetramer fraction (GCCC- and/or CCGC-cp) was discernable (*ca.* 10% within 10 weeks).

All the observations referred to above strongly point to the limits of template-controlled ligation of p-RNA 2′,3′-cyclophosphates with regard to sequence length of the activated ligands. Not surprisingly, therefore, an exploratory experiment carried out to test the behavior of a monomer-2′,3′-cyclophosphate indicated that there is hardly a chance for monomers to act as cumulative ligands in chain elongations under such conditions: no primer elongation on the G-rich dodecamer template could be observed at 4° under otherwise standard conditions in the presence of pr(C)-2′,3′-cyclophosphate (3.6 mM, 24 equiv. within 12 weeks. Whether monomer-2′,3′-cp's could insert themselves into single-unit 'ligation holes', where they would be assisted on both sides with stable preligation base pairing of oligomers remains to be seen.

8. Ligation Reactions Involving Adenine and Thymine Units at the Ligation Site. – All template-controlled ligations discussed so far use p-RNA sequences which consist exclusively of guanine- and cytosine-ribopyranosyl units. To the extent that ligation rates can be expected to depend on the pre-ligation equilibria between free template and free ligands on the one hand, and the ternary pre-ligation complexes on the other, transcription of template sequences consisting of adenine and thymine units should be slower than that of guanine-cytosine sequences of similar template and ligand length. However, such a difference should hold only for ligations in which the ligands are of critical length, *e.g.*, tetramers, but not necessarily for ligations of ligands with sequences that are long enough to form stable pre-ligation complexes under the reaction conditions. The limited set of data we have accumulated on this question qualitatively confirms this expectation.

The sequence pr(AAATAAAT) has the same purine-pyrimidine motif as our workhorse sequence pr(GGGCGGGC) and, therefore, was chosen for comparing the transcription propensity of A/T templates *vs.* G/C templates. Under conditions where the G/C analog is transcribed to an extent of *ca.* 14% within a week (1.0M LiCl, pH 8.0, room temperature), the reaction formulated in *Fig. 23* hardly takes place (less than 1% in a week). In 1.5M LiCl at pH 8.5 and 4°, it is distinctly faster, but still much slower (3% after 1, 25% after 12 weeks) than the analogous G/C ligation under same conditions (but

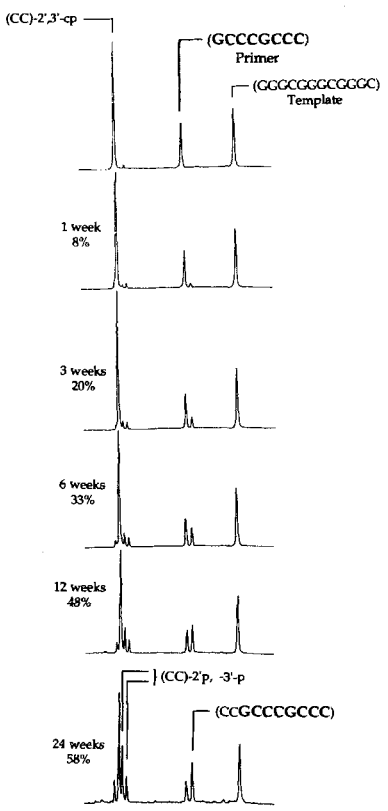


Fig. 21. *HPLC Traces of the ligation of 1.8 mM pr(CC)-2',3'-cp in the presence of 0.15 mM pr(GGGCGGGCGGGC) as template and 0.15 mM pr(GCCCGCCC) as primer* (for reaction scheme, see *Fig. 20, a*). Reaction in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, 4°.

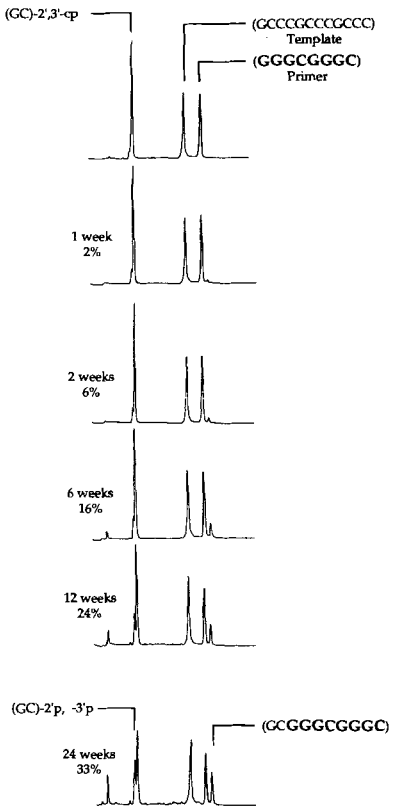


Fig. 22. *HPLC Traces of the ligation of 0.9 mM pr(GC)-2',3'-cp in the presence of 0.15 mM pr(GCCCGCCCGCCC) as template and 0.15 mM pr(GGGCGGGC) as primer* (for reaction scheme, see *Fig. 20, b*). Reaction in 1.5M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.5, 4°.

at 25°), presumably reflecting a compromise between two consequences of the lower reaction temperature, namely, a higher stability of the pre-ligation complex and a lower intrinsic ligation rate. At 4°, not only is the intrinsic ligation rate reduced, but also the hydrolytic deactivation of the cyclophosphate group; as a consequence, the reaction mixtue for pr(AAATAAAT) copying remains remarkably clean even after a long reaction time (see HPLC traces in *Fig. 23*)[21]).

We also have tested ligation for adenine-thymine units containing more than four bases. At 10° and in 1M LiCl at pH 8.0, the two octamers shown in *Fig. 24* are converted,

[21]) Unfortunately, rate data for pr(GGGCGGGC) transcription at 4° were not collected.

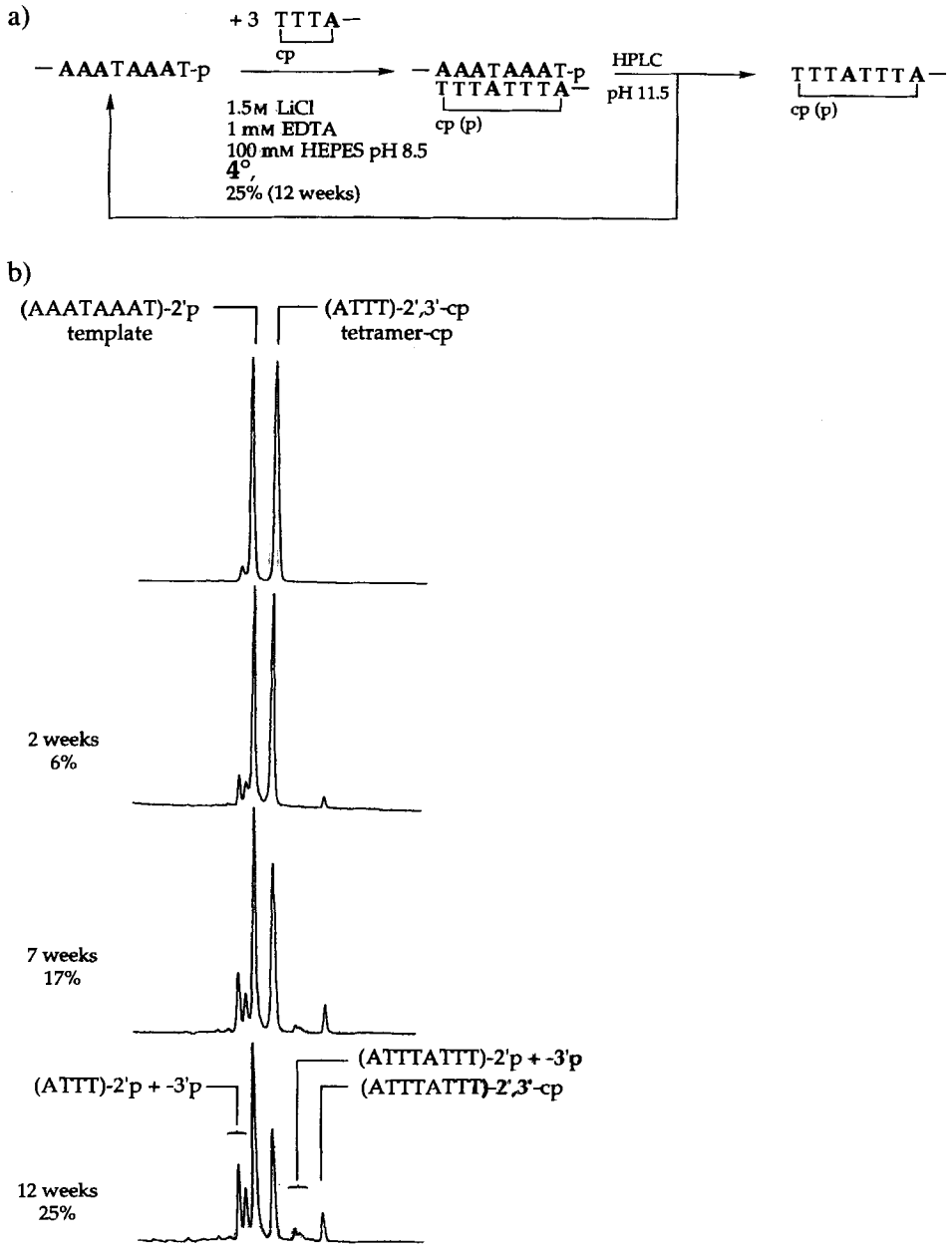


Fig. 23. *Reaction scheme* (a) *and HPLC traces* (b) *of the ligation of 0.45 mM pr(ATTT)-2',3'-cp in the presence of 0.15 mM pr(AAATAAAT)-2'-p as template*

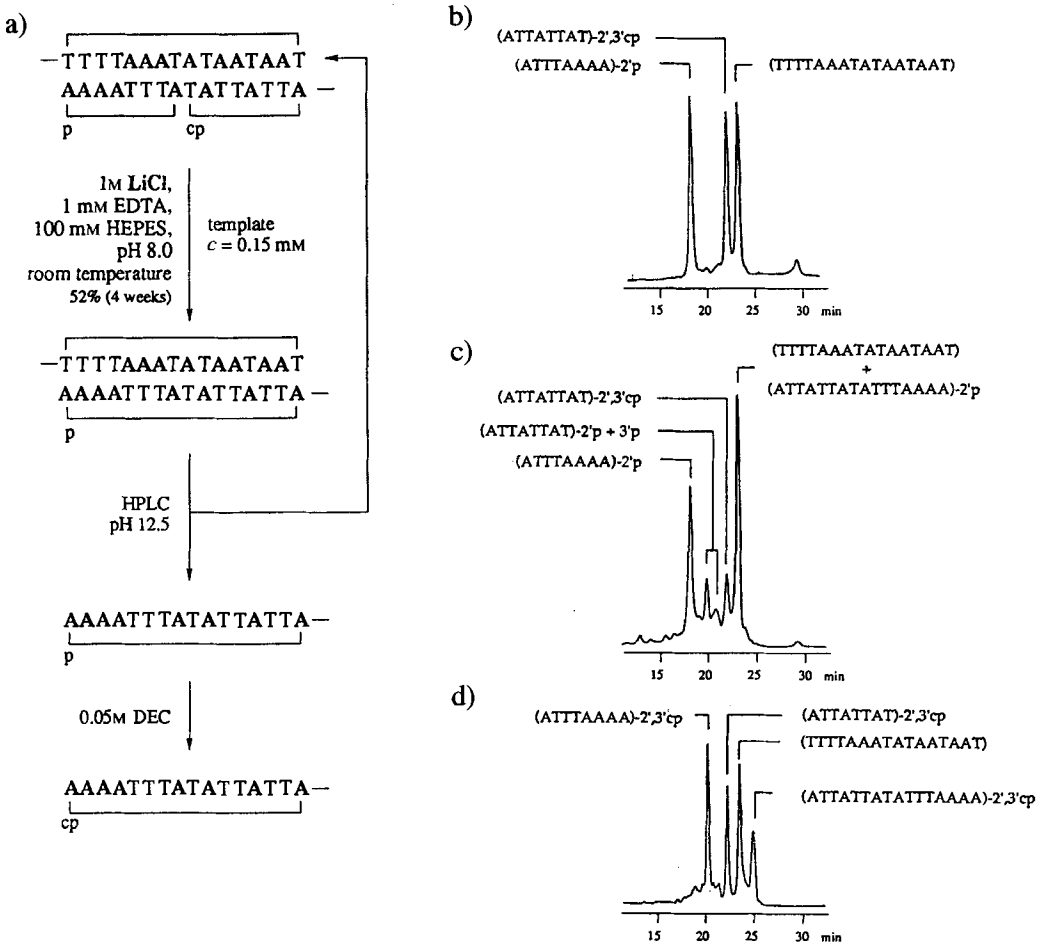


Fig. 24. *Reaction scheme* (a) *of the ligation of pr(ATTTAAAA)-2'-p and pr(ATTATTAT)-2',3'-cp* (0.15 mM each) *in the presence of 0.075 mM pr(TTTTAAATATAATAAT) as template* (1.0M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.0, 10°, HPLC Traces (pH 12.5): start (*b*) and after 3 weeks of ligation before (*c*) and after (*d*) the treatment with 0.05M DEC (2 h, r.t.) in H_2O. Yield: 45% (3 weeks), 52% (4 weeks) of pr(ATTATTATATTTAAAA)-2',3'-cp. Product identification after dephosphorylation (before DEC treatment) with alkaline phosphatase (see *Fig. 10*) by HPLC coinjection with authentic material.

under the influence of the indicated hexadecamer sequence[22]) as template, into the ligated hexadecamer at a relatively normal rate; the yield plateau (56%) is reached after 5 weeks. The product hexadecamer[22]) was identified by HPLC comparison with

[22]) This hexadecamer had been synthesized, together with the complementary hexadecamer, in another connection (comparison of constellational pairing selectivity of DNA, RNA, and p-RNA [28]). In the present study, the ligation experiment was planned in such a way that the hexadecamer sequence was generated as its 2'-phosphate. In this form it could be more easily distinguished from the template hexadecamer in HPLC analysis, especially after conversion into the 2',3'-cyclophosphate by DEC (see *Fig. 24*).

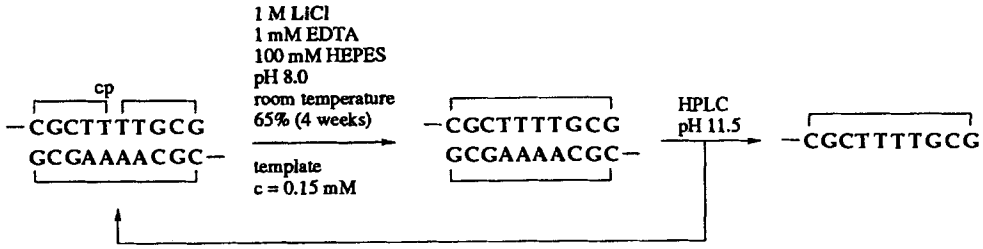


Fig. 25. *Reaction scheme of the ligation of pr(CGCTT)-2',3'-cp and pr(TTGCG)* (0.3 mM each) *in the presence of 0.15 mM pr(CGCAAAAGCG) as template* (1.0M LiCl, 1 mM EDTA, 100 mM HEPES, pH 8.0, room temperature). Yield: 65% (4 weeks) of pr(CGCTTTTGCG) [29]. Template exists as hairpin (see [5]).

an authentic sequence. Another example which shows that adenine-thymine ligation can proceed at normal rates, provided that the stability of the pre-ligation complex is high enough, is the replicative ligation of the (hairpin forming) sequence pr-(CGCAAAAGCG) (*Fig. 25*). The reaction proceeds in yields up to 70% within 3 weeks (to be discussed in more detail in a forthcoming publication [29]).

9. Qualitative Relationships between Relative Ligation Rate and Interstrand Base Stacking at Ligation Sites. – One of the important lessons which our study of the structure and reactivity of p-RNA has taught us with respect to the pairing and ligation properties of this oligonucleotide system concerns the role of *inter*strand (*vs. intra*strand) base stacking for the stability of isomeric p-RNA duplexes [18] as well as for the relative rates of ligation in template-controlled sequence copying [4] and ligative oligomerization [6]. In p-RNA, *inter*strand stacking dominates over *intra*strand stacking as a consequence of the characteristically large inclination of (approximated) backbone and base-pair axes. Relevant structural aspects have been illustrated by *Schlönvogt et al.* [1] and are reproduced here again in *Fig. 26* to assist the discussion. A recent report from our laboratory [5] describes strong evidence supporting the view that it is predominantly

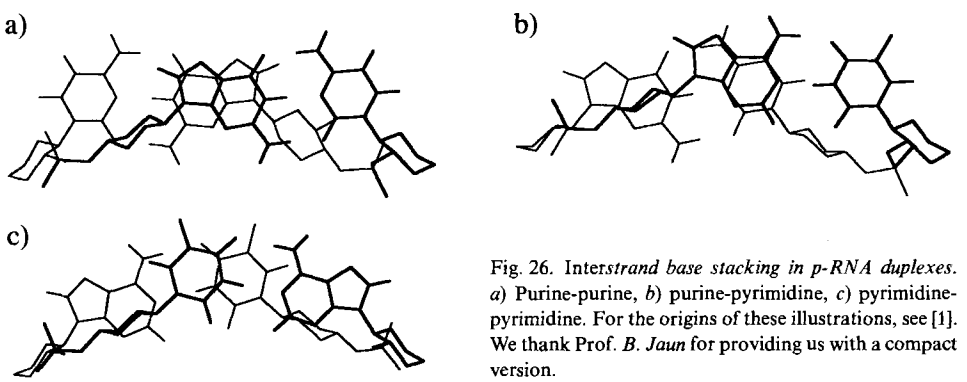


Fig. 26. Inter*strand base stacking in p-RNA duplexes. a*) Purine-purine, *b*) purine-pyrimidine, *c*) pyrimidine-pyrimidine. For the origins of these illustrations, see [1]. We thank Prof. *B. Jaun* for providing us with a compact version.

*inter*strand base-stacking that determines the highly specific interfluence of dangling bases on the stability of p-RNA duplexes. Here, we insert (partially unpublished) data on the stability of p-RNA duplexes that further illustrate the relationship between duplex stability and *inter*strand stacking; these data may serve as background for the analysis of the interdependence of ligation rates and base stacking modes at ligation sites that follows.

Table 3 summarizes melting temperatures and thermodynamic data for a series of p-RNA duplexes in which self-complementary octamer and hexamer sequences with the following sequence motifs are compared with each other (to be read in the 4′ → 2′ direction): $(py_n\text{-}PU_n)$ with $(PU_n\text{-}py_n)$, and $(py\text{-}PU)_n$ with $(PU\text{-}py)_n$. In each pair, the motif in which py (= pyrimidine) preceeds PU (= purine) forms the more stable duplex, the difference within the first motif pair being more pronounced than in the second[23]). Although the range of data provides too narrow a basis for more than a tentative interpretation, we consider its juxtaposition with the analysis given in *Fig. 27* as relevant.

Table 3. *Comparison of* T_m *Values and Thermodynamic Data of p-RNA Duplexes with Self-complementary Sequence Motifs Purine → Pyrimidine and Pyrimidine → Purine.* Data for $pr(AT)_4$ (see also [4]) are redetermined under exactly the same conditions which were used for determination of the $pr(TA)_4$ data.

	Melting temp. [a] T_m (*ca.* 10 μM)	Thermodynamic data [b] [Kcal/mol]		
		ΔG^{298}	ΔH	$T\Delta S^{298}$
pr(AAAAUUUU)	14° [18] [c])	−5.2	−42.5	−37.3
pr(UUUUAAAA)	32° [18] [c])	−8.1	−54.9	−46.8
pr(AUAUAUAU)	24° [18] [c])	−6.6	−51.1	−44.5
pr(UAUAUAUA)	27° [18] [c])	−7.2	−48.4	−41.2
pr(AAAATTTT)	27° [c])	−7.3	−48.1	−40.8
pr(TTTTAAAA)	40° [c])	−9.8	−59.9	−50.1
pr(ATATATAT)	38° [d])	−8.8	−46.6	−37.8
pr(TATATATA)	40° [d])	−9.3	−51.6	−42.3
pr(GGGCCC)	58° [e])	−10.8	−41.3	−30.4
pr(CCCGGG)	68° [e])	−13.0	−48.5	−35.5
pr(GCGCGC)	62° [4] [e])	−11.3	−40.5	−29.2
pr(CGCGCG)	65° [4] [e])	−12.5	−47.9	−35.4
$pr(A_8) + pr(U_8)$	28° [18] [c])	−8.0	−50.1	−42.1
$pr(A_8) + pr(T_8)$	40° [c])	−10.5	−62.2	−51.7
$pr(G_6) + pr(C_6)$	61° [4] [f])	−13.5	−54.3	−40.8

[a]) in 0.15M NaCl; 0.01M *Tris* · HCl; pH 7.0 [b]) Determined from concentration dependence of T_m values according to [54]. [c]) Data determined by *S. Pitsch* [55]. [d]) Data determined by *N. Windhab* [56]. [e]) Data determined by *M. Minton* [57]. [f]) Data determined by *S. Wendeborn* [58].

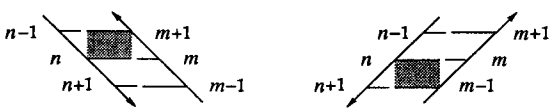
[23]) Note also the consistently higher stability of the thymine- *vs.* uracil-containing duplexes. The data for the three duplexes containing non-self-complementary strands are given for reference purposes.

The four models show that, in each case, the number of *inter*strand stacking relationships in duplexes containing the motifs (py_n-PU_n) and (py-PU)$_n$ exceeds the number of similar stackings in the duplexes with the motifs (PU_n-py_n) and (PU-py)$_n$, respectively. Within both motif pairs, the number of stackings parallels duplex stability[24]). Thereby, the pair that exhibits a more pronounced difference in duplex stabilities also can be interpreted more reliably: whereas in (py_n-PU_n) duplexes all purines are engaged in *inter*strand stacking, such is not the case in (PU_n-py_n) duplexes, the net difference between the two being a purine-purine *vs.* pyrimidine-pyrimidine *inter*strand stacking relationship in the center of the molecules. For the rest, both duplexes show the same number of purine-pyrimidine stackings. In the motif pair of duplexes with alternating bases, *inter*strand stacking is exclusively purine-purine. Here, the difference of four *vs.* three interstrand (PU-PU) stackings would lead us to expect that the difference in duplex stability should be comparable to that observed in the former pairs. This is not the case; the differences are notoriously small (see *Table 3*). Interestingly enough, if the structure of the duplex with the (PU-py)$_n$ motif would correspond to the model with the frameshift as indicated in *Fig. 27* (below, to the far right), the number of (PU-PU) interstrand stackings would become the same as in the duplex with the (py-PU)$_n$ motif. The structures of these duplexes call for a more detailed investigation[24]).

Within the four isomeric sequences of *Exper. A–D* (*Fig. 16*) ligation efficiencies differ by roughly a factor of three, as judged from the product yields after 4 weeks under

---

[24]) If in an oligonucleotide system the inclination between (approximated) backbone and base-pair axes is not positive (base *n* of strand A pairs with base *m* of strand B and stacks with base *m* + 1 of strand B [5]) as in p-RNA, but negative (base *n* stacks with base *m* – 1; see *Fig.* below), *inter*strand base stacking would dominate in duplexes with the motifs (PU_n-py_n) and (PU-py)$_n$, respectively, and not *vice versa*, as in p-RNA. The expectations on relative duplex stabilities would have to be inverse. A system with negative inclination is homo-DNA [7] [30] [25]. (For an NMR analysis of a homo-DNA duplex, see [31]). However, the (scalar) degree of inclination is less than in p-RNA (for a quantification of back-bone-base pair inclination, see a forthcoming publication by *Egli et al.* [28]). Since the importance of the inclination parameter for the interpretation of certain properties of oligonucleotide systems had not been fully realized at the time of our homo-DNA studies, and became evident only later in our studies on p-RNA, experimental data on relative duplex stabilities in the homo-DNA series, which would be useful for a comparative analysis of *inter*strand stacking, are scarce. Unfortunately, available data for comparing duplexes with (py_n-PU_n) and (PU_n-py_n) motifs are incomplete [30]. They are available for the (less informative) pair ddGlc(GC)$_3$ (T_m = 64° at 50 μM, ΔG^{298} = – 10.3) and ddGlc(CG)$_3$ (T_m = 62° at 35 μM, ΔG^{298} = – 10.6) [30], but these data are too close to allow a reliable conclusion. On the other hand, evidence in favor of the *inter*strand stacking model in homo-DNA exists in data for two duplexes that contain self-complementary strands with dangling bases and are paired in the reverse-*Hoogsteen* mode: [ddGlc(G_6-T_2)]$_2$, T_m = 39° (20 μM), ΔG^{298} = – 8.6 kcal/mol; and [ddGlc(T_2-G_6)]$_2$, T_m = 48° (20 μM), ΔG^{298} = – 0.6 kcal/mol [30]. Higher duplex stability when dangling bases are at the 6′-end (corresponding to p-RNA's 4′-end) is the opposite of what is observed with p-RNA [5]. An extension of this analysis to RNA and DNA will be discussed in [28].



POSITIVE inclination (*e.g.* p-RNA)

NEGATIVE inclination (*e.g.* homo-DNA)

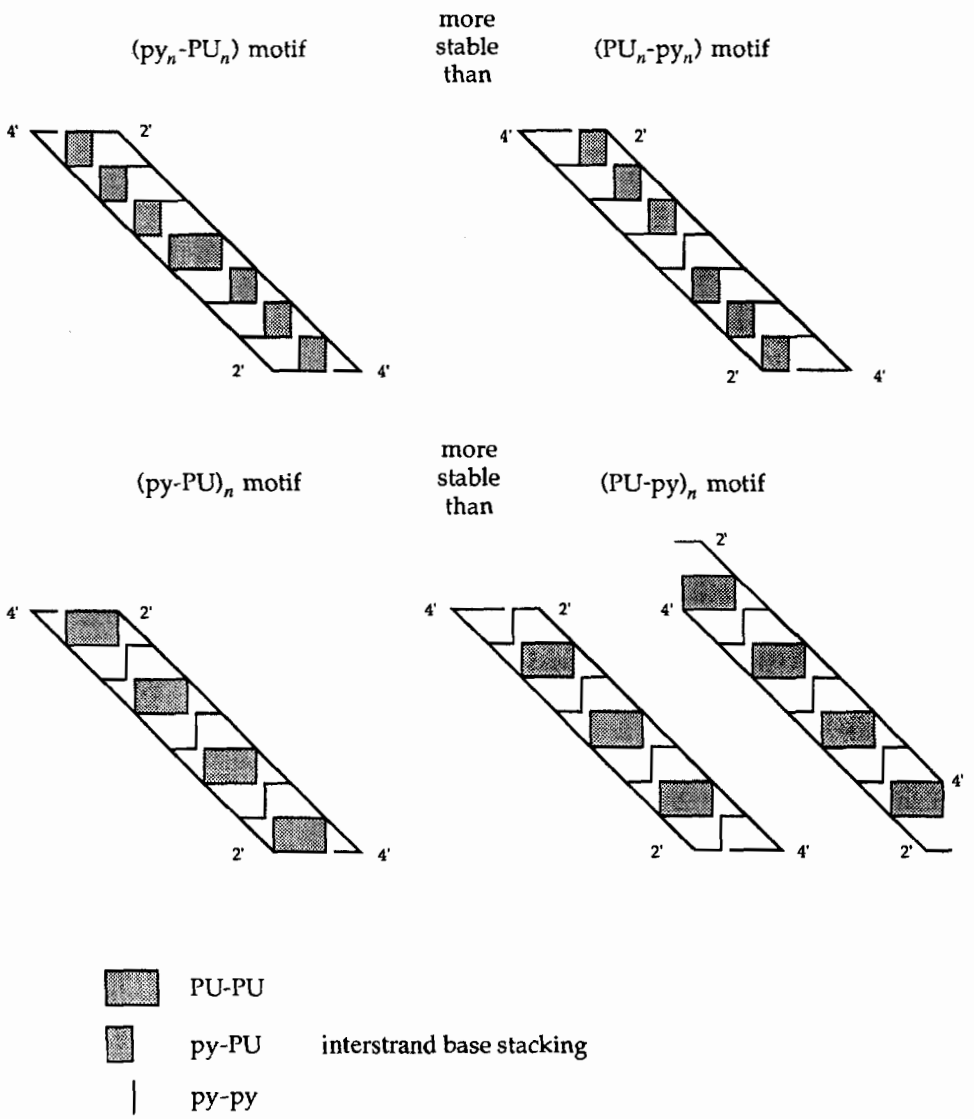


Fig. 27. *Correlation of differences in stability of p-RNA duplexes* (see *Table 3*) *with differences in* inter*strand stacking.*

identical conditions[25]). They are illustrated again in *Fig. 28* using presentations that emphasize the *inter*strand stacking differences within this series of four template ligand combinations. Three correlations come to mind: first, *inter*strand stacking at ligation sites parallels ligation efficiency, if one assumes that PU-PU stacking is more favorable

[25]) Note that these experiments were run under the less than optimal conditions, in the presence of $MgCl_2$ (see *Table 1*), and, therefore, produced low yields. Their approximate yield ratios can be expected to be valid also for the corresponding ligations in LiCl solution.

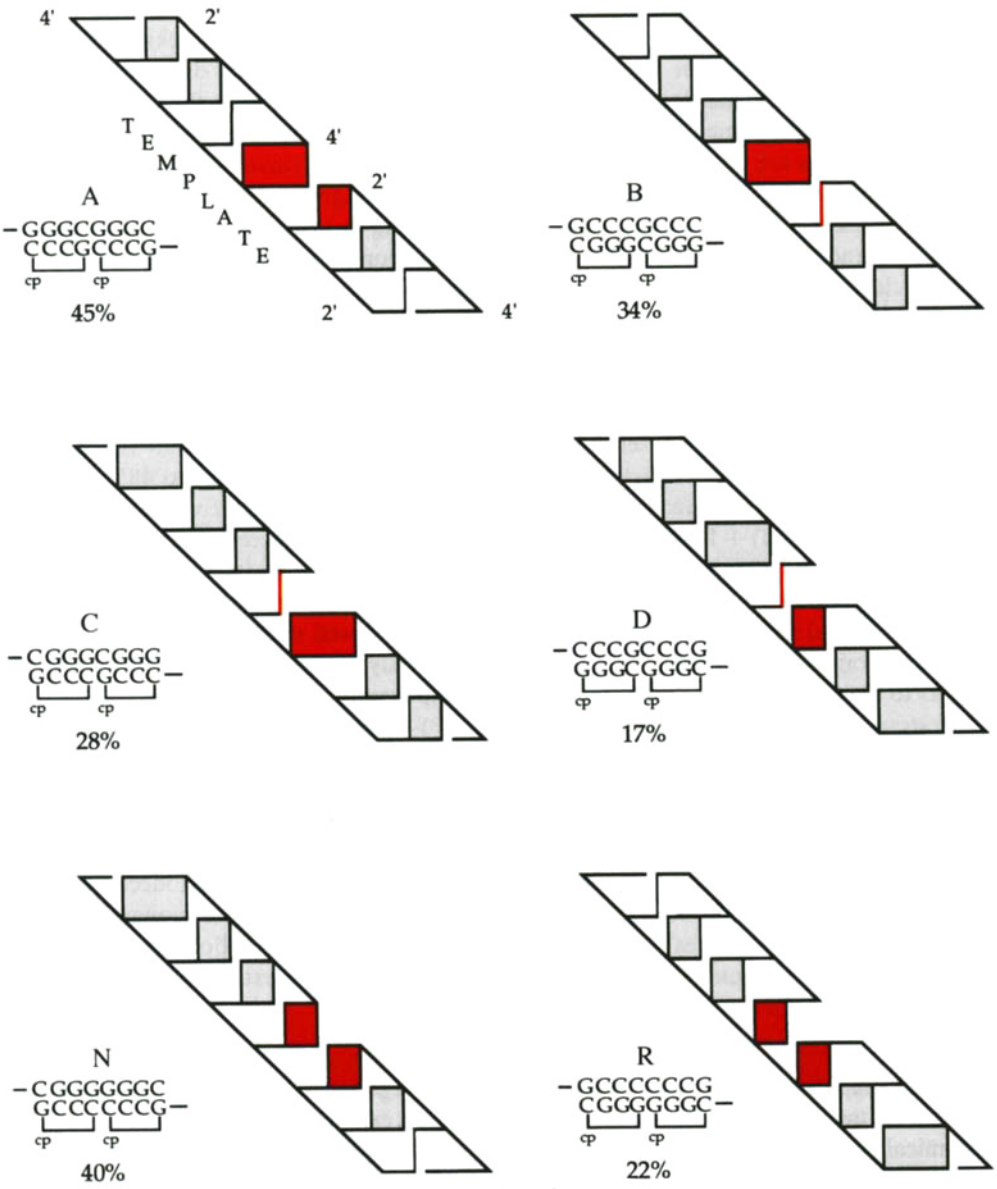


Fig. 28. *Correlation of differences in ligation rates* (see *Fig. 16*) *with differences in* inter*strand stacking at ligation sites*

than PU-py stacking; second, ligation efficiency parallels the number of template purines that stack with ligand bases at ligation sites; and, third, copying PU-rich templates is more efficient than copying complementary py-rich templates. All these correlations seem to be expressions of the same control factor, most probably conformational fixation of reaction centers for ligation through *inter*strand stacking. However, PU-py stacking

may not simply equate pyrimidine-purine stacking (see *Exper. N* and *R*): PU(template)-py(ligand) stacking could be more favorable than the inverse type of stacking. All these considerations suffer from uncertainty about the extent to which differences in the stability of the template-ligand pre-ligation complexes co-determine the observed efficiences. All six experiments involve isomeric sequence combinations so that the pre-ligation complexes are composed of the same number of G·C *Watson-Crick* pairs. Differences in the stability of these complexes would have to originate from overall differences in base stacking. A comparison of the overall stacking in the pre-ligation complexes of *Exper. A* and *D* (5 *vs.* 6 stackings) with the corresponding ligation efficiences (45 *vs.* 17%) clearly speaks against the possibility that pre-ligation complex stability is the dominant factor for the relative efficiences of ligation observed, at least in this series[26]).

The tentative conclusions drawn from these considerations are nicely corroborated by a corresponding analysis (*Fig. 29*) of the observations made in *Exper. I*, *K*, *L*, and *M* (*Fig. 16*). Among these, the ligation of *Exper. L* is the most efficient and, moreover, behaves exceptionally in the sense that in no other case does ligation go beyond the octamer copy (the potential octamer yield in *Exper. L* would amount to 48%). *Inter*strand stacking at (octamer-forming) ligation sites is optimal for *Exper. L* (PU(template) → PU(ligand) + PU(template) → py(ligand)), it is mediocre for *Exper. I* and *M* (yields 29 and 32%), and it is minimal for *Exper. K* (yield 5%; py(template) → py(ligand) + py(template → PU(ligand)). With respect to *inter*strand stacking at the potential ligation site for single-base-overlap-induced dodecamer formation, the sequence combination of *Exper. L* is unique: it is the only case in which the octamer duplex offers to the third tetramer unit a single-base-overlap ligation site at which there is effective *inter*strand stacking (PU(template) → PU(ligand)) between template and incoming tetramer; therefore, it should not be taken as a surprise that in none of the other cases is dodecamer formation observed. *Exper. L* demonstrates that single-base-overlap ligation is feasible, provided that pre-complexation is assisted by PU-PU *inter*strand stacking.

Fig. 30 explains, by related arguments, why the octamer primer in the experiments formulated in *Fig. 20* can be expected to sit at the 4′-ends of the two dodecamer templates. Furthermore, *Fig. 30, c*, makes it understandable why primer elongation with dimer-2′,3′-cyclophosphates at the guanine-rich dodecamer template does not go beyond the decamer. The argument also holds for the (slower) primer extension referred to in *Fig. 30, d*.

10. Comparison with RNA: Template-Controlled Ligation of Tetramer-2′,3′-cyclophosphates in the RNA Series. – Fundamental to our experimental approach toward a chemical etiology of nucleic-acid structure is the concept of systematically comparing, on a chemical level, relevant properties of nucleic-acid alternatives with the corresponding

---

[26]) Based on observations made in an early phase of the ligation studies, we did not attempt to determine relative stabilities of the pre-ligation complexes of *Exper. A–D*, and *N* and *R*. The early observations had been: 4 μM $pr(G_8)$ in the presence of 8 μM $pr(C_4)$-4′-phosphate (in 0.15M NaCl, pH 7.0) produced a UV melting curve that clearly indicated complexation below *ca.* 40° (*ca.* 10% hyperchromicity) but was not sigmoidal between 0°–40°. A CD spectroscopically (but not the UV spectroscopically) monitored mixing curve determined at 0° indicated the formation of both a 1:1, as well as a 1:2 complex (5 μM). The T_m value of the duplex $pr(G_8) \cdot pr(C_8)$ (2.5 + 2.5 μM) under the same conditions is 82° (10 + 10 μM) [4].

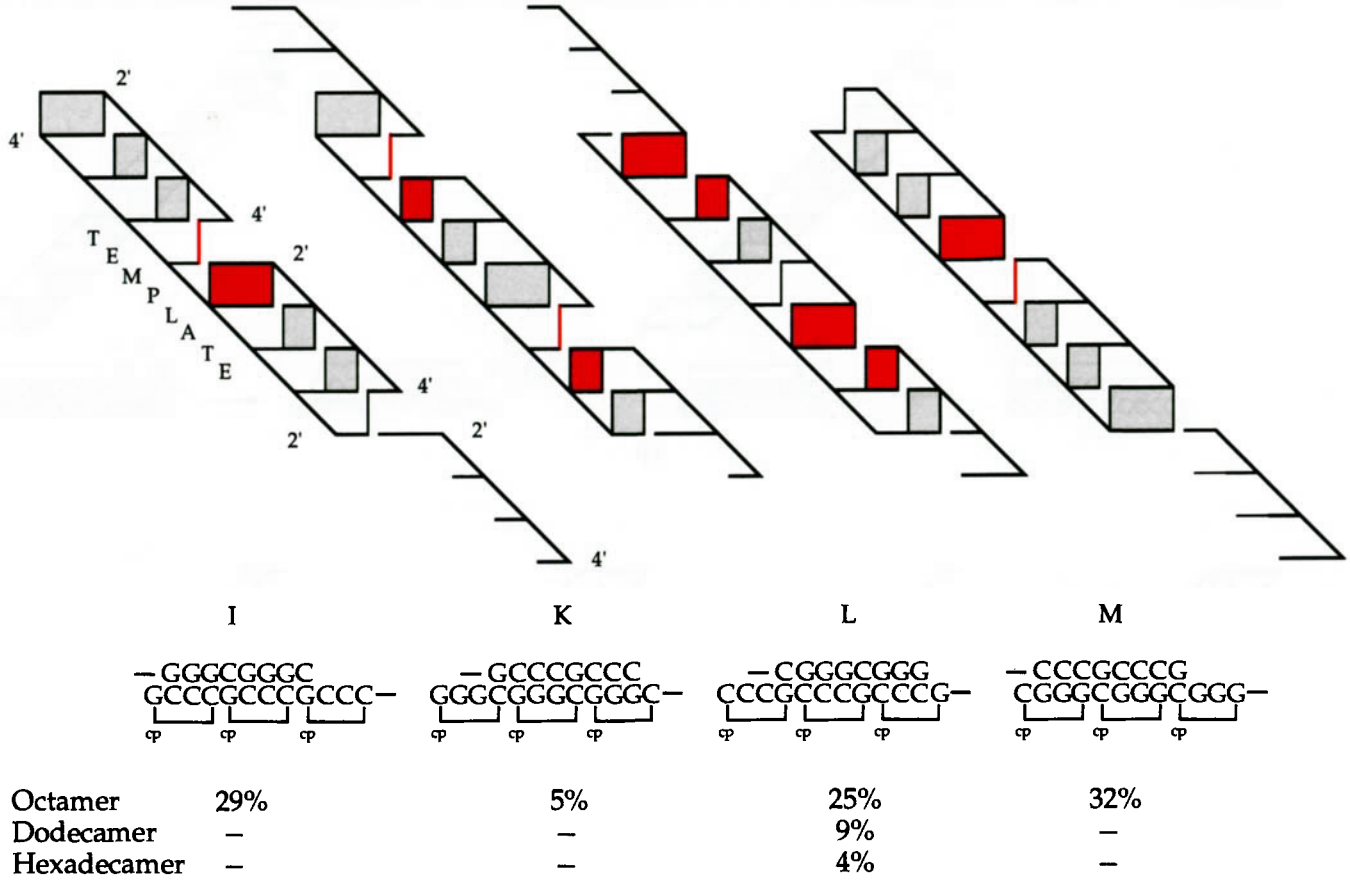


Fig. 29. *Correlation of differences in ligation rates and oligomerization propensites* (see *Fig. 16*) *with differences in* inter*strand stacking at ligation sites*

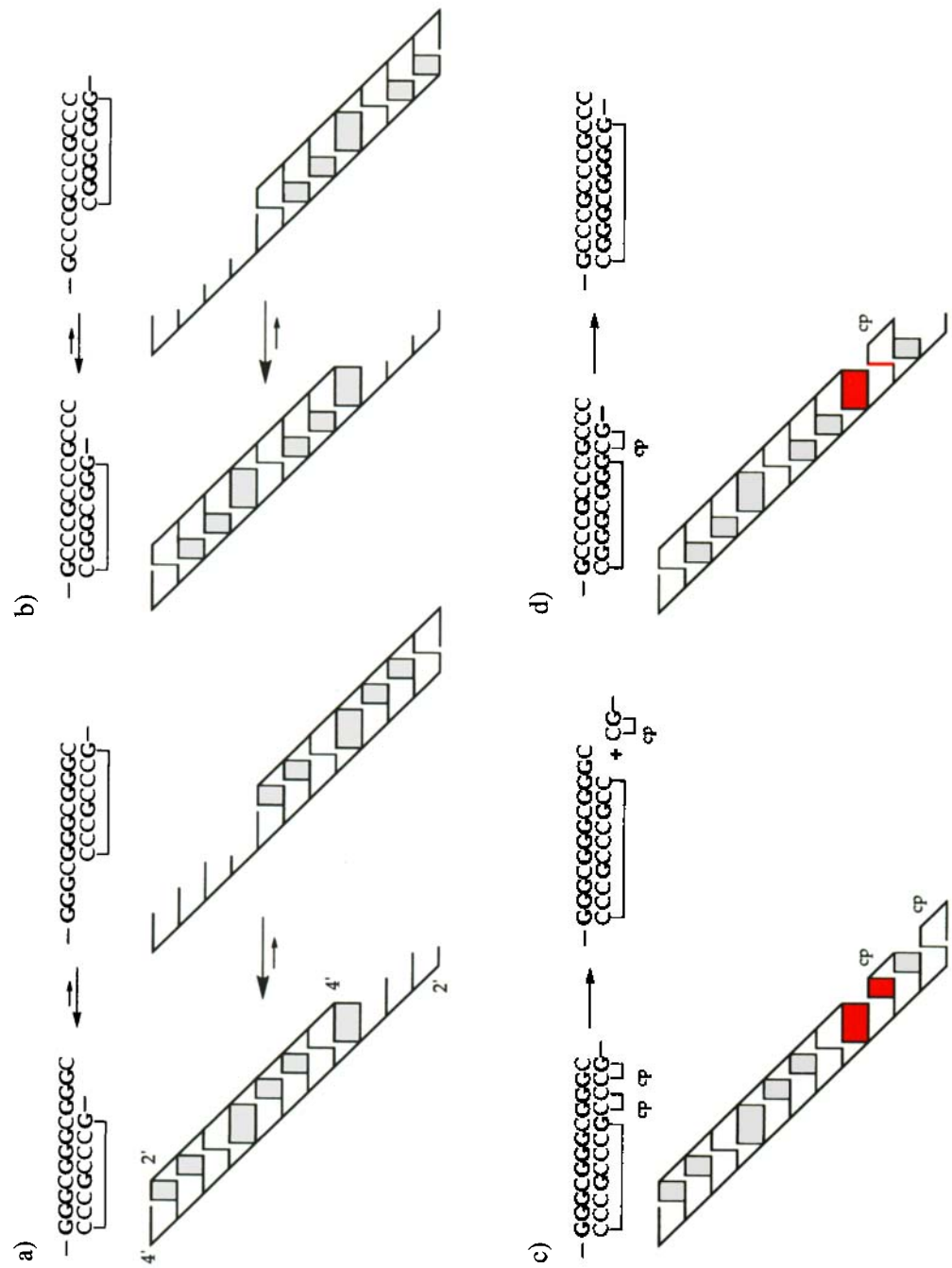


Fig. 30. *Constitutional assignment for the template primer duplexes* (*a* and *b*) *and interpretation of the course of primer elongation with dimer-2′,3′-cyclophosphates* (*c* and *d*) *by* interstrand *base-stacking analysis* (see *Figs. 20–22*)

(chemical) properties of RNA [24] [32–34]. With this in mind, we have investigated the behavior of the sequence $^{5'}$GGGCGGGC$^{3'}$ with regard to its capacity for replicative ligation of the corresponding tetramer-2′,3′-cyclophosphates in the RNA series under conditions similar to those used in the p-RNA series.

In an early paper on template-controlled oligomerization of activated mononucleotides in the RNA series, *Orgel* and co-workers [11] remarked '*since all modes of activation of ribofuranosyl nucleoside 2′(3′)-phosphates in aqueous solution led to the formation of 2′,3′-cyclophosphates, these substances stand out as candidates for prebiotic activated intermediates*'. Experimentally, he observed adenosine-2′,3′-cyclophosphate in 0.025M solution at pH 8.0 on a poly(U) template in the presence of diamines as catalysts, preferably ethylenediamine, to produce the nucleotide dimer (together with the mixture of corresponding 2′- and 3′-phosphates as hydrolysis products) in yields up to 23%. Hydrolysis of the cyclophosphate group was, under all tested conditions, faster than transesterification to the dimer, and, presumably for this reason, only traces of the corresponding trimer were observed. The phosphodiester junction in the dimer turned out to be predominantly 2′ → 5′. In a follow-up study, *Orgel* and co-workers [35] [36] also showed that adenosine-2′,3′-cyclophosphate self-oligomerized to oligomer chains of up to at least six units by being kept in a 'dry state' at moderate temperatures and again in the presence of diamines as catalysts – the mode of phosphodiester junctions now being, in contrast to the template-controlled reaction referred to above – almost two times as much 3′ → 5′ as 2′ → 5′. Apparently as a consequence of the shortcomings of these reactions, the ligation chemistry of ribonucleotide-2′,3′-cyclophosphates in the furanosyl-RNA series was not pursued further[27]), with the exception of the ligation experiment of *Usher* and *McHale* [12], in which the self-ligation of r(A_6)-2′,3′-cyclophosphate on a poly(U) template under *Orgel*'s conditions [11] (1M ethylenediamine, pH 8.0, 2°, 5 d) was shown to lead in 24% yield to a corresponding dodecamer that contained a central r(A_6-A_6) bond determined to be predominantly 2′ → 5′ (*Fig. 31*).

It was specifically our guanine-rich sequence GGGCGGGC that we chose for testing the 2′,3′-cyclophosphate-mediated ligative replication in the RNA series for comparing p-RNA and RNA. A determinant of this choice was, not least, the recurring difficulties encountered by *Orgel* and co-workers [8] [10] [37] [38] in attempting transcriptions of guanine-rich sequences by template-controlled oligomerization of *mono*nucleotide-5′-phosphoimidazolides into corresponding cytosine-rich sequences, due to template inactivation through self-pairing of the guanine-rich sequences. Whether this type of copying handicap of the RNA series would persist, when ligands were not activated mononucleotides, but activated tetramer units, was one of the uncertainties to be settled experimentally. We also wished to address the question of how 2′,3′-cyclophosphates in the RNA series would behave under conditions where ligative replication in the p-RNA series proceeds smoothly, and, if they were to ligate, whether under these conditions they would lead to (2′ → 5′) or (3′ → 5′) phosphodiester junctions (*Fig. 31*).

---

[27]) '... *we think it unlikely that they* (the cyclic 2′,3′-phosphates) *were the substrate of the first template-catalyzed polymerization, that led to the formation of long oligonucleotides.* ...,', p. 375 in [10].

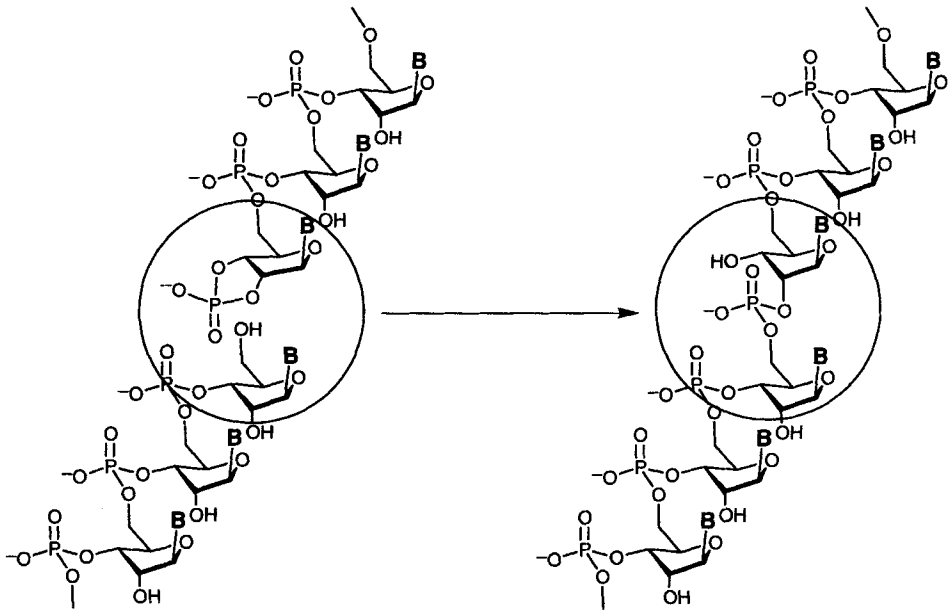


Fig. 31. *Comparing pyranosyl-RNA with furanosyl-RNA with respect to template-controlled ligation of tetramer-2',3'-cyclophosphates.* Does the reaction under the conditions used in the p-RNA series also proceed in the RNA series, and, if yes, does it lead to (3' → 5') or (2' → 5') phosphodiester junctions?

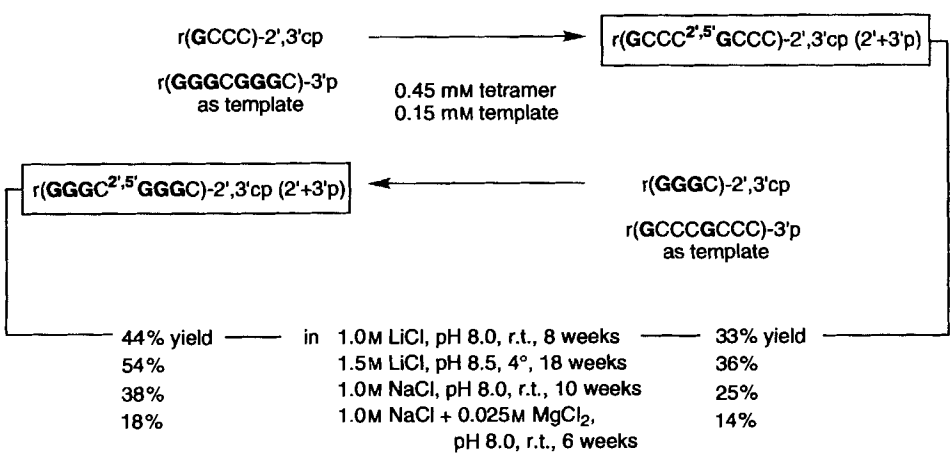


Fig. 32. *Template-directed ligation of r(GCCC)-2',3'-cp* (0.45 mM) *with r(GGGCGGGC)-3'-p* (0.15 mM) *as template and of r(GGGC)-2',3'-cp* (0.45 mM) *with r(GCCCGCCC)-3'-p* (0.15 mM) *as template under a variety of conditions* (all reactions in the L-ribofuranosyl series). Both reactions produce exclusively the (2' → 5') phosphodiester junction between the fourth and fifth ribofuranosyl units of the product octamers (see *Fig. 34*). Yields calculated with respect to template and determined by HPLC (same conditions as indicated in *Fig. 4*).

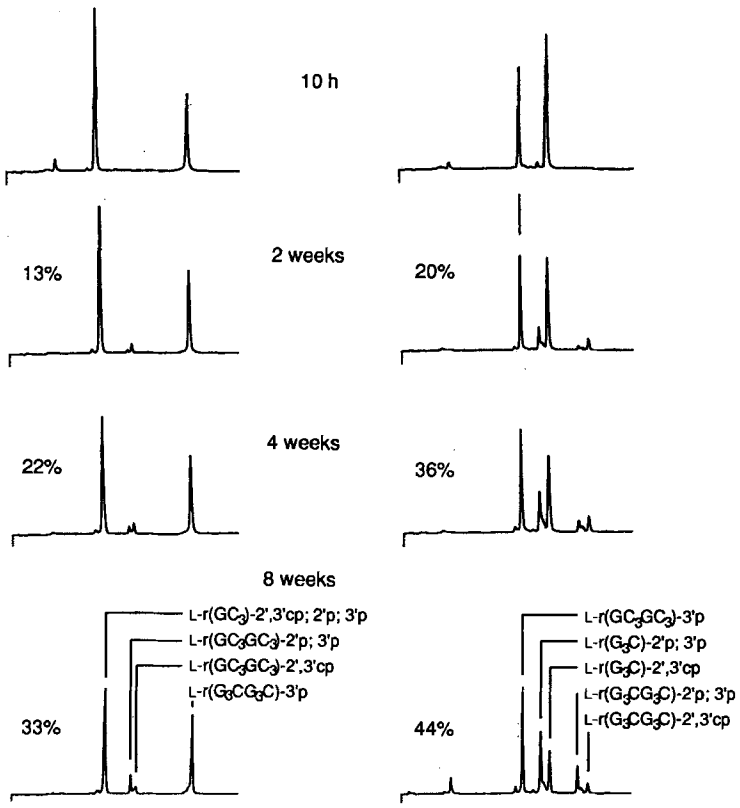


Fig. 33. *HPLC Traces of the ligation of 0.45 mM* L-*r(GCCC)-2',3'-cp* (left) *and 0.45 mM* L-*r(GGGC)-2',3'-cp* (right), *each in the presence of 0.15 mM of indicated templates in 1.0M LiCl, pH 8.0, room temperature.* HPLC Conditions see *Fig. 4.* No strand scission observed under reaction conditions. No octamer formation observed in the absence of templates (< 0.5%). For constitutional assignment of product peaks, see *Fig. 34.*

The ligation experiments which were carried out (all of them in the unnatural (nuclease resistant) L-series[28]) are formulated in *Fig. 32*, and the results are summarized and documented in *Figs. 33* and *34*.

Transcription of the guanine-rich sequence r(GGGCGGGC) (as its 3'-phosphate) under a variety of conditions (see *Fig. 32*) proceeds *ca.* 5 times more slowly than the corresponding process in p-RNA, with a ligation/hydrolysis ratio that is comparable, yet somewhat less in favor of ligation than it is in the pyranosyl series. It is also slower than the transcription of the complementary cytosine-rich sequence r(GCCCGCCC) under the same conditions (*Fig. 33*, right side), showing a relationship of transcription efficiencies between the two templates opposite to that observed in the pyranosyl series. The rates of the latter transcription are similar in the furanosyl and pyranosyl series. The

[28]) The L-series was chosen because one of us (*S.P.*), in context with an independent project, recently has worked out a simplified method for the synthesis of RNA-oligonucleotides in the L-series [39].

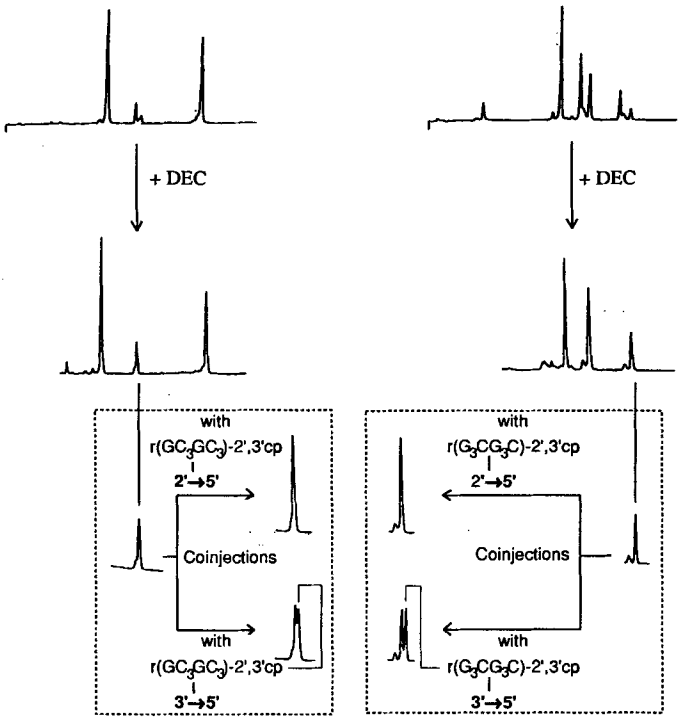


Fig. 34. *Identification of products of ligation reaction referred to in* Fig. 33 *by HPLC co-injection of products with authentic samples of the octamer-2',3'-cyclophosphates containing a (2' → 5') phosphodiester junction, between the forth and the fifth ribofuranosyl unit and the corresponding (3' → 5')-isomers, respectively.* The data show no discernible amounts of (3' → 5')-isomers. For the synthesis of the authentic product isomers, see [39].

dependence of both ligations on reaction conditions (*Fig. 32*) seems, more or less, to parallel the corresponding dependence observed in the pyranosyl series.

The two ligation reactions investigated in the RNA series proceed – as they do in the pyranosyl series – with high regioselectivity: the template-controlled ligation of the two tetramer-2',3'-cyclophosphates produces exclusively the (2' → 5') phosphodiester junction. This is demonstrated unambiguously by the experiments documented in *Fig. 34*. Both octamer product fractions (mixtures of 2',3'-cyclophosphates and 2'- and 3'-phosphates) were treated with water-soluble carbodiimide in order to convert the phosphate components into the cyclophosphate form, and the product cyclophosphate fractions were analyzed by HPLC co-injections with authentic fractions[29]). The peaks of the ligation products sharply coincided with the peaks of (synthetic) r(GCCC$^{2',5'}$GCCC)-2',3'-cp and r(GGGC$^{2',5'}$GGGC)-2',3'-cp, respectively, but were unambiguously different in retention time from the peaks of authentic r(GCCC$^{3',5'}$GCCC)-2',3'-cp and

[29]) These comparisons make clear why the template sequences were synthesized as their 3'-phosphates, rather than in their 3'-OH form, as in the earlier experiments in the pyranosyl series. For the synthesis of the octamer sequences containing the central (3' → 5') and (2' → 5') phosphodiester junctions, see the forthcoming publication by *Pitsch* [39].

r(GGGC$^{3',5'}$GGGC)-2',3'-cp, respectively (*Fig. 34*). The shapes of the observed HPLC traces of the two ligation products indicated the absence of any isomeric components with a retention time that would correspond to that of the (3' → 5')-isomers (regioselectivity estimated to be higher than 98%).

This outcome of our comparison of the pyranosyl and furanosyl series with respect to the course of template-controlled replicative ligation of tetramer 2',3'-cyclophosphates points to the following conclusion: apart from a trend toward lower efficiency in the furanosyl series, the behavior of the two systems is similar. At least for templates as short as octamers, there seems to be no serious barrier to transcribing a guanine-rich sequence into the corresponding cytosine-rich sequence by ligation of tetramer-2',3'-cyclophosphates in the RNA series. In view of the difficulties observed in attempts to replicate G-rich RNA sequences by template-controlled oligomerization of activated monomers [38], this observation is important, though not necessarily surprising. Naturally, a more extended investigation using a variety of template sequences would be desirable in order to decide how similarly G-rich RNA and p-RNA template sequences behave in such ligations.

What has been adumbrated by the earlier observations of *Orgel* and co-workers [11] and *Usher* [12] is corroborated by our findings: in the furanosyl series, template-controlled replicative ligation of ligands as their 2',3'-cyclophosphates leads to (2' → 5') phosphodiester junctions with high regioselectivity. Mechanistically, this parallels the regioselective formation of (2' → 4') phosphodiester junctions in the pyranosyl series, and the constellational as well as stereoelectronic reasons for both regioselectivities are presumably analogous. However, this close chemical relationship stands in sharp contrast to the etiological perspective: template-controlled ligation *via* 2',3'-cyclophosphates produced the 'right' phosphodiester junction in pyranosyl-RNA, but the 'wrong' junction in RNA. 'Right' and 'wrong' refer here to a consideration, according to which in p-RNA the (2' → 4') system corresponds to the 'natural' member among the two possible isomers, in as far as it is the one with the optimal base-pairing capability. On the other hand, in the furanosyl series the (2' → 5') system is known to be far inferior to natural RNA as a base-pairing system [40].

11. Concluding Remarks. – What *Orgel* and co-workers [11] in their early work explicitly considered for the RNA series, but later abandoned [10] as a result of the outcome of exploratory experiments [35] [36], may eventually regain validity in the p-RNA series: ribonucleotide 2',3'-cyclophosphates could indeed be etiologically relevant. However, we should not necessarily adhere to the postulate [41] according to which non-enzymatic replication in oligonucleotide systems must be based on the principle of one-by-one mononucleotide transcription. While such a transcription mode clearly has the potential for the highest possible degree of replication fidelity, at the same time it is the chemically most demanding mode, since it must operate with minimum base-pairing and base-stacking stabilization in the template-ligand pre-ligation complex. Pre-organization of ligands is more easily attained with (activated) short oligomers as compared to monomers. Considering the differences in overall rate of ligation due to differences in pre-organization equilibria, a non-enzymatic template-driven oligomerization of activated mononucleotides in the absence of specific primer sequences may in any case proliferate into a complex series of processes which includes template-controlled

ligation of short oligonucleotides of varying lengths [10]. In retrospect, it seems not surprising that the first example of non-enzymatic informational replication of hexameric DNA-oligonucleotide derivatives by *von Kiedrowski* [14] was based on template-controlled ligation of trimers.

In a research strategy of searching systematically for the potential for non-enzymatic replication of nucleic-acid alternatives in comparison to that of RNA [24] [32], it seems quite realistic to concentrate first on the capability to replicate by template-controlled ligation of short oligonucleotides, instead of insisting on one-by-one mononucleotide copying. The latter, especially including turnover, may be considered to represent an evolved level of achievement that may require specific catalysis [38].

In the etiological context, it is tempting to emphasize the virtues of p-RNA's 2′,3′-cyclophosphate group as a potentially natural form of phosphate activation and to set it against the strengths and shortcomings of 5′-phosphate activation modes[30]) that have been observed to lead to (5′ → 3′) phosphodiester junctions in template-driven oligomerizations of mononucleotides in the RNA series. 2′,3′-Cyclophosphate groups offer a simple, robust, and mild form of phosphate activation in a ribonucleotide system. In p-RNA, they have the remarkable and highly welcome property of undergoing virtually no ligative reactions in dilution, except under the influence of an oligonucleotide template. As products of an intramolecular dehydration, they are easily and reliably accessible, both in mononucleotides, or at 2′(or 3′)-ends of corresponding oligomers [6]. Sensitivity toward hydrolysis is moderate and, above all, the outcome of hydrolysis can, in principle as well as in practice [6], be reversed at any stage of an oligomerization process by intramolecular dehydration. Stoichiometrically, cyclophosphate-mediated ligations are isomerizations: The products of ligation are isomeric with the ensemble of ligands. In principle, oligonucleotide sequences and ligand cyclophosphates could be in equilibrium with each other; therefore, sequence libraries could be formed under *thermodynamic* control, and specific duplexes could select themselves from such libraries according to the criterion of duplex stability. If oligoribonucleotide systems as such are considered to have been prebiologically relevant, then 2′,3′-cyclophosphates should be considered to have been so as well.

In view of our observations on 2′,3′-cyclophosphates properties in the p-RNA series and the findings discussed in *Chapt. 9* regarding the regioselective formation of (2′ → 5′) phosphodiester junctions by template-controlled ligation in the RNA series, it becomes of increasing interest to include the (2′ → 5′)-ribofuranosyl system in the systematic comparison of nucleic-acid alternatives with natural RNA, even though it has been reported to be a pairing system far inferior to RNA, at least with respect to pairing strength

---

[30]) The strength of *Orgel*'s 5′-phosphoimidazolide derivatives [10] [41] is their propensity to efficiently mediate template-induced oligocondensation of monomers. This capability is mainly a consequence of their high electrophilic reactivity and betainic (neutral) nature in their *N*(3)-protonated form [10]. The high-powered activation has its shortcomings in that it can lead to template-unrelated background reactivity [42] [43] and implies a high sensitivity toward hydrolysis. 5′-Phosphoimidazolides derived from oligonucleotide-5′-triphosphates recently have been shown by *Bartel* and *Szostak*, and *Rohatgi et al.* [44–46] to mediate non-enzymatic template-directed ligation of oligoribonucleotides and to lead regioselectively to (3′ → 5′)-phosphodiester junctions. Due to extensive deactivation by hydrolysis, the ligation rate relative to the rate with triphosphate activation could not be quantified but was estimated to be faster by a factor of at least 500 [46]. Ligation of triphosphates themselves on RNA templates is extremely slow ($t_{1/2}$ = 15–30 years, pH 7.4; 100 mM Mg^{++}) [45].

[40]. Comparison of the two systems was pioneered by *Ts'o* and [47] and *Usher* [12] in their attempts to rationalize the existence of the (3′ → 5′) junction as opposed to the (2′ → 5′) junction in RNA[31]). The formation of (2′ → 5′)-isomers in addition to (3′ → 5′)-RNA oligomers has been a recurringly observed phenomenon in *Orgel*'s [10] [8], and more recently, in *Ferris*' [50] studies on the oligomerization of activated ribofuranosyl mononucleotides. Our own work on p-RNA raises the specific question whether (2′ → 5′)-ribofuranosyl template sequences also could be copied by replicative ligation of 2′,3′-cyclophosphates of short (2′ → 5′)-RNA oligomers.

Whatever level of achievement in non-enzymatic copying of sequence information in oligonucleotide systems eventually may be reached, the barrier that separates such capabilities from the potential of oligonucleotide systems to evolve as families of molecules is product inhibition in the transcription process[32]). To discover, or to design, solutions by which a system would overcome this barrier in a general way, must be a major goal of synthetic informational chemistry focusing on oligonucleotide systems. It is in this context that the progress of ongoing research aiming at enzymatic *in vitro* evolution of RNA sequences that would be able to act as non-proteinic RNA polymerases [53] is followed with special interest.

This work was supported by *Ciba-Geigy AG*, Basel, and the *Schulleitung of ETH*. *A.E.* would like to express his personal thanks to Dr. *D. Bellus*, Dr. *G. Haas*, Dr. *W. Fuhrer*, Dr. *P. Schneider*, and Dr. *H. Moser* (*Ciba-Geigy AG*), and to Prof. Dr. *R. Hütter* (*ETH*). The work could not have been carried out without the administrative support of the Laboratory of Organic Chemistry at ETH. *R.M.* thanks the *Fonds zur Förderung der wissenschaftlichen Forschung*, Wien, for a *Schrödinger* postdoctoral fellowship. We thank Dr. *R.M. Wolf* (*Novartis AG*, Basel) and Prof. *B. Jaun* (ETH) for kindly providing us with *Fig. 12* and *26*, respectively. *A.E.* thanks Prof. *Claude Wintner*, Haverford College, PA (USA), for checking the English.

Experimental Part

General. Experiments dealing with the preparation of p-RNA sequences will be described in [17]. An overview of all ligation experiments referred to in this paper is given in *Table 4*.

HPLC: *Pharmacia GP-250* gradient programmer equipped with two *P-500* pumps, *ABI-Kratos-Spectroflow-757* UV/VIS detector and a *Hewlett Packard HP 3396A* integrator. Column: *Mono Q HR 5/5* (*Pharmacia*). Matrix-assisted laser-desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) of ligation products was carried out in the laboratory of Dr. *U. Pieles*, *Ciba-Geigy AG* (see [23]). UV Melting curves were determined on a *Perkin Elmer Lambda 2* spectrometer, equipped with a *Perkin Elmer* digital controller/temp. programmer *C570*; the temp. gradient was ~ 0.8°/min; the temp. was measured directly in the sample soln.

---

[31]) More recently, the groups of both *Breslow* and co-workers [48] and *Switzer* and co-workers [49] have studied rather systematically the chemistry of the 3′-deoxyribofuranosyl-(2′ → 5′)-oligonucleotide system in comparison to DNA. Whereas the results of these studies are of great interest with regard to the relationships between base-pairing and backbone structure, it should be noted that the (2′ → 5′)-DNA system, quite in contrast to (2′ → 5′)-RNA, is only indirectly of etiological relevance.

[32]) Informational replication with turnover has been observed by *von Kiedrowski* [14] [51] in ligations of (structurally modified) DNA trimers catalyzed by the complementary hexamers with water-soluble carbodiimide as coupling reagent. So far, similar behavior with longer template sequences has not been reported. Informational replication with turnover of polypeptide sequences recently has been achieved by *Ghadiri* and co-workers [52].

Table 4. *Overview of Ligation Experiments.* Conditions: *a*: 1.0M NaCl, 25 mM $MgCl_2$, 100 mM HEPES pH 8.0, r.t.; *b*: 1.5M LiCl, 1 mM EDTA, 100 mM HEPES pH 8.5, r.t.; *c*: 1.5M LiCl, 1 mM EDTA, 100 mM HEPES pH 8.5, 4°; *d*: 1.0M LiCl, 1 mM EDTA, 100 mM HEPES pH 8.0, 10°; *e*: same as *d*, 4°; *f*: same as *d*, r.t.; *g*: 1.0M NaCl, 1 mM EDTA, 100 mM HEPES pH 8.0, r.t. Yields with respect to template. *Exper. 39–46* carried out with ribofuranosyl sequences of the L-series.

Exper.	Template (150 μM)	Cyclophosphate(s) (450 μM)	Reaction conditions	Ligation products	Yield [%] (after ... weeks)	Exper. data documented in
1	(G_3CG_3C)	(GC_3)-2′,3′-cp	*a*	(GC_3GC_3)-2′,3′-cp + (GC_3GC_3))-2′/3′-p	23 (1 w), 37 (2 w), 45 (3 w), 45 (4 w)	[4]
2	(G_3CG_3C)	(GC_3)-2′,3′-cp	*b*	(GC_3GC_3))-2′,3′-cp + (GC_3GC_3))-2′/3′-p	57 (1 w), 57 (2 w)	*Fig. 4, c*; *5, a*; *7*; *9*; *10*; *13, a*; *15, a*
3	(G_3CG_3C)	L-(GC_3)-2′,3′-cp	*b*	–	0 (1–2 w)	*Fig. 13, b*
4	(G_3CG_3C)	L-(GC_3)-2′,3′-cp + D-(GC_3)-2′,3′-cp (450 μM each)	*b*	D-(GC_3GC_3))-2′,3′-cp + D-(GC_3GC_3)-2′/3′-p	58 (1 w), 60 (2 w)	*Fig. 13, c*; *14, d–e*
5	(G_3CG_3C)	(GCC-L-C)-2′,3′-cp	*b*	–	0 (1–2 w)	*Fig. 15, b*
6	(G_3CG_3C)	(GC-(L-C)-C)-2′,3′-cp	*b*	–	0 (1–2 w)	*Fig. 15, c*
7	(G_3CG_3C)	(G-L-C)-CC)-2′,3′-cp	*b*	–	0% (1–2 w)	*Fig. 15, d*
8	(G_3CG_3C)	((L-G)-CCC)-2′,3′-cp	*b*	((L-G)-CCC((L)-C)-CCC)-2′,3′-cp + ((L-G)-CCC((L)-G)-CCC)-2′/3′-p	11 (1 w), 14 (2 w)	*Fig. 15, e*
9	(GC_3GC_3)	(G_3C)-2′,3′-cp	*a*	(G_3CG_3C)-2′,3′-cp + (G_3CG_3C)-2′/3′-p	20 (1 w), 27 (2 w), 34 (4 w)	
10	(GC_3GC_3)	(G_3C)-2′,3′-cp	*b*	(G_3CG_3C)-2′,3′-cp + (G_3CG_3C)-2′/3′-p	52 (1 w), 54 (2 w)	*Fig. 6, a*; *7*; *9*
11	(CG_3CG_3)	C_3G)-2′,3′-cp	*a*	(C_3GC_3G)-2′,3′-cp + (C_3GC_3G)-2′/3′-p	17 (1 w), 25 (2 w), 28 (3.5 w), 28 (4 w)	*Fig. 35, C*

Table 4 (cont.)

Exper.	Template (150 μM)	Cyclophosphate(s) (450 μM)	Reaction conditions	Ligation products	Yield [%] (after ... weeks)	Exper. data documented in
12	(C_3GC_3G)	(CG_3)-2′,3′-cp	*a*	(CG_3CG_3)-2′,3′-cp + (CG_3CG_3)-2′/3′-p	11 (1 w), 15 (2 w), 17 (4 w)	*Fig. 35, D*
13	(G_3CG_3C)	(G_3C)-2′,3′-cp	*a*	–	0 (1–6 w)	
14	(GC_3GC_3)	(GC_3)-2′,3′-cp	*a*	–	0 (1–6 w)	
15	(CG_3CG_3)	(CG_3)-2′,3′-cp	*a*	–	0 (1–6 w)	*Fig. 35, G*
16	(C_3GC_3G)	(C_3G)-2′,3′-cp	*a*	–	0 (1–6 w)	*Fig. 35, H*
17	(G_3CG_3C)	(C_3G)-2′,3′-cp	*a*	(C_3GC_3G)-2′,3′-cp + (C_3GC_3G)-2′/3′-p	14 (1 w), 23 (2 w), 29 (3 w)	
18	(GC_3GC_3)	(CG_3)-2′,3′-cp	*a*	(CG_3CG_3)-2′,3′-cp + (CG_3CG_3)-2′/3′-p	4 (1 w), 4 (2 w), 5 (3 w), 5 (6 w)	
19	(CG_3CG_3)	GC_3)-2′,3′-cp	*a*	(GC_3GC_3)-2′,3′-cp + (GC_3GC_3)-2′/3′-p	17 (1 w), 19 (2 w), 28 (3 w), 38 (4 w)	*Fig. 35, L*
20	(C_3GC_3G)	(G_3C)-2′,3′-cp	*a*	(G_3CG_3C)-2′,3′-cp + (G_3CG_3C)-2′/3′-p	19 (1 w), 29 (2 w), 35 (3 w), 35 (4 w),	
21	(CG_6C)	(GC_3)-2′,3′-cp + (C_3G)-2′,3′-cp (225 μM each)	*a*	(GC_6G)-2′,3′-cp + (GC_6G)-2′/3′-p	40 (4 w)	*Fig. 35, N*
22	(CG_6C)	(GC_3)-2′,3′-cp + (C_4)-2′,3′-cp (225 μM each)	*a*	(GC_7)-2′,3′-cp + (GC_7)-2′/3′-p	33 (4 w)	

Table 4 (cont.)

Exper.	Template (150 μM)	Cyclophosphate(s) (450 μM)	Reaction conditions	Ligation products	Yield [%] (after ... weeks)	Exper. data documented in
23	(CG_6C)	(C_3G)-2′,3′-cp + (C_4)-2′,3′-cp (225 μM each)	*a*	(C_7G)-2′,3′-cp + (C_7G)-2′/3′-p	22 (4 w)	
24	(CG_6C)	(C_3G)-2′,3′-cp	*a*	–	0 (1–4 w)	
25	(GC_6G)	(CG_3)-2′,3′-cp + (G_3C)-2′,3′-cp (225 μM each)	*a*	(CG_6C)-2′,3′-cp + (CG_6C)-2′/3′-p	13 (1 w), 18 (2 w), 22 (3 w)	*Fig. 35, R*
26	(GC_6G)	(CG_3)-2′,3′-cp	*a*	–	0 (1–3 w)	
27	(GC_6G)	(G_3C)-2′,3′-cp	*a*	(G_3CG_3C)-2′,3′-cp + (G_3CG_3C)-2′/3′-p	1 (1 w), 2 (2 w), 2 (3 w)	
28	(A_3TA_3T)-2′p	(AT_3)-2′,3′-cp	*c*	(AT_3AT_3)-2′,3′-cp + (AT_3AT_3)-2′/3′-p	3 (1 w), 6 (2 w), 17 (7 w), 25 (12 w)	*Fig. 23, b*
29	($G_3CG_3CG_3C$)	(GC_3)-2′,3′-cp (1.35 mM)	*a*	(GC_3GC_3)-2′,3′-cp + (GC_3GC_3)-2′/3′-p + ($GC_3GC_3GC_3$)-2′,3′-cp + ($GC_3GC_3GC_3$)-2′/3′p	30 8-mers, 26 12-mers (4 w)	
30	($G_3CG_3CG_3C$)	(GC_3)-2′,3′-cp (1.35 mM)	*b*	(GC_3GC_3)-2′,3′-cp + (GC_3GC_3)-2′/3′-p + ($GC_3GC_3GC_3$)-2′,3′-cp + ($GC_3GC_3GC_3$)-2′/3′p	28 8-mers, 56 12-mers (2 w)	*Fig. 18, b*
31	($G_3CG_3CG_3C$)	(GC)-2′,3′-cp + (CC)-2′,3′-cp (2.7 mM each)	*c*	(GCCC)-2′,3′-cp + (GCCC)-2′/3′-p and/or (CCGC)-2′,3′-cp + (CCGC)-2′/3′-p	2 (1 w), 2 (2 w), 4 (3 w), 5 (4 w), 9 (10 w)	

Table 4 (cont.)

Exper.	Template (150 μM)	Cyclophosphate(s) (450 μM)	Reaction conditions	Ligation products	Yield [%] (after ... weeks)	Exper. data documented in
32	($G_3CG_3CG_3C$)	(GC_3GC_3) (150 μM) + (GC)-2′,3′-cp + (CC)-2′,3′-cp (0.9 mM each)	*c*	($CCGC_3GC_3$)	6 (1 w), 11 (2 w), 20 (4 w), 33 (8 w)	*Fig. 21*
33	($G_3CG_3CG_3C$)	(GC_3GC_3) (150 μM) + (CC)-2′,3′-cp (1.8 mM)	*c*	($CCGC_3GC_3$)	8 (1 w), 15 (2 w), 33 (6 w), 48 (12 w), 58 (24 w)	*Fig. 22*
34	($G_3CG_3CG_3C$)	(GC_3GC_3) (150 μM + (C)-2′,3′-cp (3.6 mM)	*c*	–	0 (1–12 w)	
35	($GC_3GC_3GC_3$)	(G_3C)-2′,3′-cp (1.35 mM)	*a*	(G_3CG_3C)-2′,3′-cp + (G_3CG_3C)-2′/3′-p + ($G_3CG_3CG_3C$)-2′,3′-cp + ($G_3CG_3CG_3C$)-2′/3′-p	33 8-mers, 10 12-mers (4 w)	
36	($GC_3GC_3GC_3$)	(G_3C)-2′,3′-cp (1.35 mM)	*b*	(G_3CG_3C)-2′,3′-cp + (G_3CG_3C)-2′/3′-p + ($G_3CG_3CG_3C$)-2′,3′-cp + ($G_3CG_3CG_3C$)-2′/3′-p	36 8-mers, 32 12-mers (2 w)	*Fig. 19, b*
37	($T_4A_3TATA_2TA_2T$) (75 μM)	(AT_2AT_2AT)-2′,3′-cp + (AT_3A_4)-2′-p (150 μM each)	*d*	($AT_2AT_2ATAT_3A_4$)-2′-p	19 (1 w)[a]) 35 (2 w)[a]), 45 (3 w)[a]), 52 (4 w)[a])	*Fig. 24, b–d*
38	($CGCA_4GCG$)	(CGCTT)-2′,3′-cp + (TTGCG) (300 μM each)	*e*	($CGCT_4GCG$)	37 (1 w), 54 (2 w), 62 (3 w), 65 (4 w)	[29]
39	r(G_3CG_3C)	r(GC_3)-2′,3′-cp	*a*	r($GC_3^{2',5'}GC_3$)-2′,3′-cp + r($GC_3^{2',5'}GC_3$)-2′/3′-p	14 (6 w)	

Table 4 (cont.)

Exper.	Template (150 μM)	Cyclophosphate(s) (450 μM)	Reaction conditions	Ligation products	Yield [%] (after ... weeks)	Exper. data documented in
40	r(G_3CG_3C)	r(GC_3)-2′,3′-cp	*c*	r($GC_3^{2'5'}GC_3$)-2′,3′-cp + r($GC_3^{2'5'}GC_3$)-2′/3′-p	36 (18 w) 8 (1 w)	
41	r(G_3CG_3C)	r(GC_3)-2′,3′-cp	*f*	r($GC_3^{2',5'}GC_3$)-2′,3′-cp + r($GC_3^{2',5'}GC_3$)-2′/3′-p	13 (2 w), 22 (4 w), 27 (6 w), 33 (8 w)	*Fig. 33*
42	r(G_3CG_3C)	r(GC_3)-2′,3′-cp	*g*	r($GC_3^{2',5'}GC_3$)-2′,3′-cp + r($GC_3^{2',5'}GC_3$)-2′/3′-p	25 (10 w)	
43	r(GC_3GC_3)	r(G_3C)-2′,3′-cp	*a*	r($G_3C^{2',5'}G_3C$)-2′,3′-cp + r($G_3C^{2',5'}G_3C$)-2′/3′-p	18 (6 w)	
44	r(GC_3GC_3)	r(G_3C)-2′,3′-cp	*c*	r($G_3C^{2'5'}G_3C$)-2′,3′-cp + r($G_3C^{2'5'}G_3C$)-2′/3′-p	54 (18 w)	
45	r(GC_3GC_3)	r(G_3C)-2′,3′-cp	*f*	r($G_3C^{2'5'}G_3C$)-2′,3′-cp + r($G_3C^{2'5'}G_3C$)-2′/3′-p	14 (1 w), 21 (2 w), 36 (4 w), 43 (6 w), 44 (8 w)	*Fig. 33*
46	r(GC_3GC_3)	r(G_3C)-2′,3′-cp	*g*	r($G_3C^{2'5'}G_3C$)-2′,3′-cp + r($G_3C^{2'5'}G_3C$)-2′/3′-p	38 (10 w)	

[a]) Yield determined after conversion of ($AT_2AT_2ATAT_3A_4$)-2′-p into the corresponding cyclophosphate with DEC.

CD Spectra were collected on a *JASCO J-710* spectropolarimeter. Extinction coefficients of p-RNA sequences were calculated as the sum of the corresponding monomer extinctions (*Table 5*).

Table 5. *Extinction Coefficients of p-RNA Monomers*

	pH 7.0		pH 11.5	
	$\varepsilon_{260\ nm}$	$\varepsilon_{270\ nm}$	$\varepsilon_{260\ nm}$	$\varepsilon_{270\ nm}$
pr(A)	15000	11200	15100	11500
pr(T)	10000	9800	10400	9600
pr(C)	8400	9100	8500	9300
pr(G)	11900	9900	12000	11500

Ligation Experiments. For each of the oligonucleotides to be used in ligation experiments a stock soln. of 0.5 to 1.2 optical densities (OD_{270}) was prepared. The exact concentration was determined by dividing the measured UV absorbance at 270 nm by the extinction coefficient of the given p-RNA sequence. Unless not stated otherwise, the concentration of template in the ligation experiments was 150 μM, whereas the concentration of tetramer cyclophosphate was 450 μM. The corresponding volume of stock soln. containing the amount of oligonucleotide needed for the set-up was pipeted into an *Eppendorf* tube. After the addition of all the sequence needed for the particular ligation experiment, the mixture was lyophilized to dryness. Finally, to the dry material the appropriate ligation buffer was added. The soln. was vortexed and briefly centrifuged. The reaction was kept either at r.t. (22–25°) or at 4–6°, according to the particular set-up. For ligation experiments which were only followed by HPLC, the reaction volume was 10–25 μl (*i.e.*, 6–15 nmol of oligonucleotide). In those cases where the ligation products were isolated and characterized by HPLC, TOF MS, UV melting curves, and CD spectra, the reaction volume was 400–500 μl (*i.e.*, 240–300 nmol of oligonucleotide).

Prior to taking the samples for HPLC analysis, the reaction soln. was vortexed and briefly centrifuged. In general, samples of 2 μl were taken. The samples were diluted in 400 μl of H_2O and directly injected onto the HPLC system without any further treatment. All analyses were performed on a *Mono Q HR 5/5* column (*Pharmacia*) with a linear gradient of 0 to 1M NaCl in 10 mM Na_2HPO_4 at pH 11.5, detection at 270 nm.

Ligation yields were determined in two different ways. In the case of the ligation of D-pr(GCCC)-2′,3′-cp, the yield of ligation products was determined by adding a known amount of D-pr(GCCCGCCC) as external standard to an aliquot of the ligation reaction mixture. This mixture was desalted over a *Sep-Pak®* cartridge and analyzed by HPLC. The procedure was repeated three times, each time adding a different amount of the external standard to an aliquot of the mixture. The ratio of the peak areas of ligation products plus external standard to template is linearly depending on the amount of standard added. Hence, the yield of ligation products was extrapolated by linear regression of the two parameters. An analogous procedure was followed to determine the yield of the ligation reaction of D-pr(GGGC)-2′,3′-cp. Here, D-pr(GGGCGGGC) was used as external standard. In all other cases, the ligation yield was calculated as the relative amount of ligation products with respect to the sum of all tetramer cyclophosphate derivatives (*i.e.*, tetramer phosphates, tetramer cyclophosphate, and ligation products) multiplied by 1.5 (the tetramer cyclophosphate is used in a 1.5-fold excess with respect to the template). The ligation yield of the reaction of D-pr(GCCC)-2′,3′-cp in the presence of D-(GGGCGGGC) was also determined according to the second method. In general, the two results obtained by the two methods were within the range of accuracy of the methods (± 2% yield), the yields determined according to the second method being somewhat higher.

Constitutional Assignment of Ligation Products. In the case of the ligation of pr(GCCC)-2′,3′-cp in the presence of pr(GGGCGGGC) as well as of the ligation of pr(GGGC)-2′,3′-cp in the presence of pr(GCCCGCCC), the ligation products were separated by HPLC. The identity of the isolated products was confirmed by MALDI-TOF MS (*Table 2*), coinjection with authentic material in HPLC (*Fig. 10*), UV melting behavior (*Fig. 9*) as well as CD spectra (*Fig. 8*). In all other cases, the assignment of the ligation products is based on their observed HPLC retention time compared to the expected retention time. This assignments are further supported by the observation of the typical pattern in the HPLC trace generated by the ligation product phosphate and its corresponding cyclophosphate (two peaks separated by *ca.* 2 min in t_R). In addition, the observation of the relative amounts of products during the course of the reaction served as another criterion for peak assignment. On the one hand, the amount of product cyclophosphate first increases due to ligation, but decreases during the course of the reaction

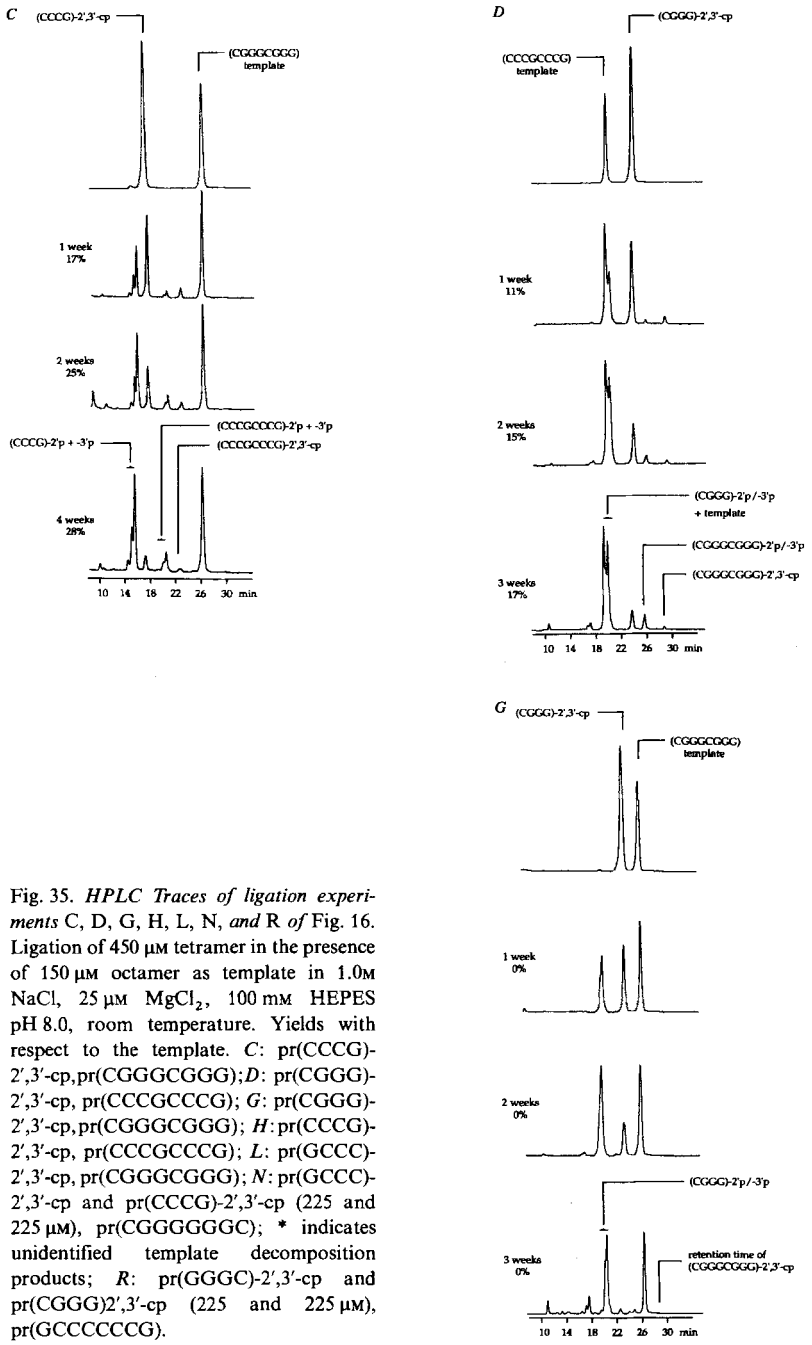


Fig. 35. *HPLC Traces of ligation experiments* C, D, G, H, L, N, *and* R *of* Fig. 16. Ligation of 450 μM tetramer in the presence of 150 μM octamer as template in 1.0M NaCl, 25 μM $MgCl_2$, 100 mM HEPES pH 8.0, room temperature. Yields with respect to the template. *C*: pr(CCCG)-2′,3′-cp,pr(CGGGCGGG);*D*: pr(CGGG)-2′,3′-cp, pr(CCCGCCCG); *G*: pr(CGGG)-2′,3′-cp,pr(CGGGCGGG); *H*:pr(CCCG)-2′,3′-cp, pr(CCCGCCCG); *L*: pr(GCCC)-2′,3′-cp, pr(CGGGCGGG); *N*: pr(GCCC)-2′,3′-cp and pr(CCCG)-2′,3′-cp (225 and 225 μM), pr(CGGGGGGC); * indicates unidentified template decomposition products; *R*: pr(GGGC)-2′,3′-cp and pr(CGGG)2′,3′-cp (225 and 225 μM), pr(GCCCCCCG).

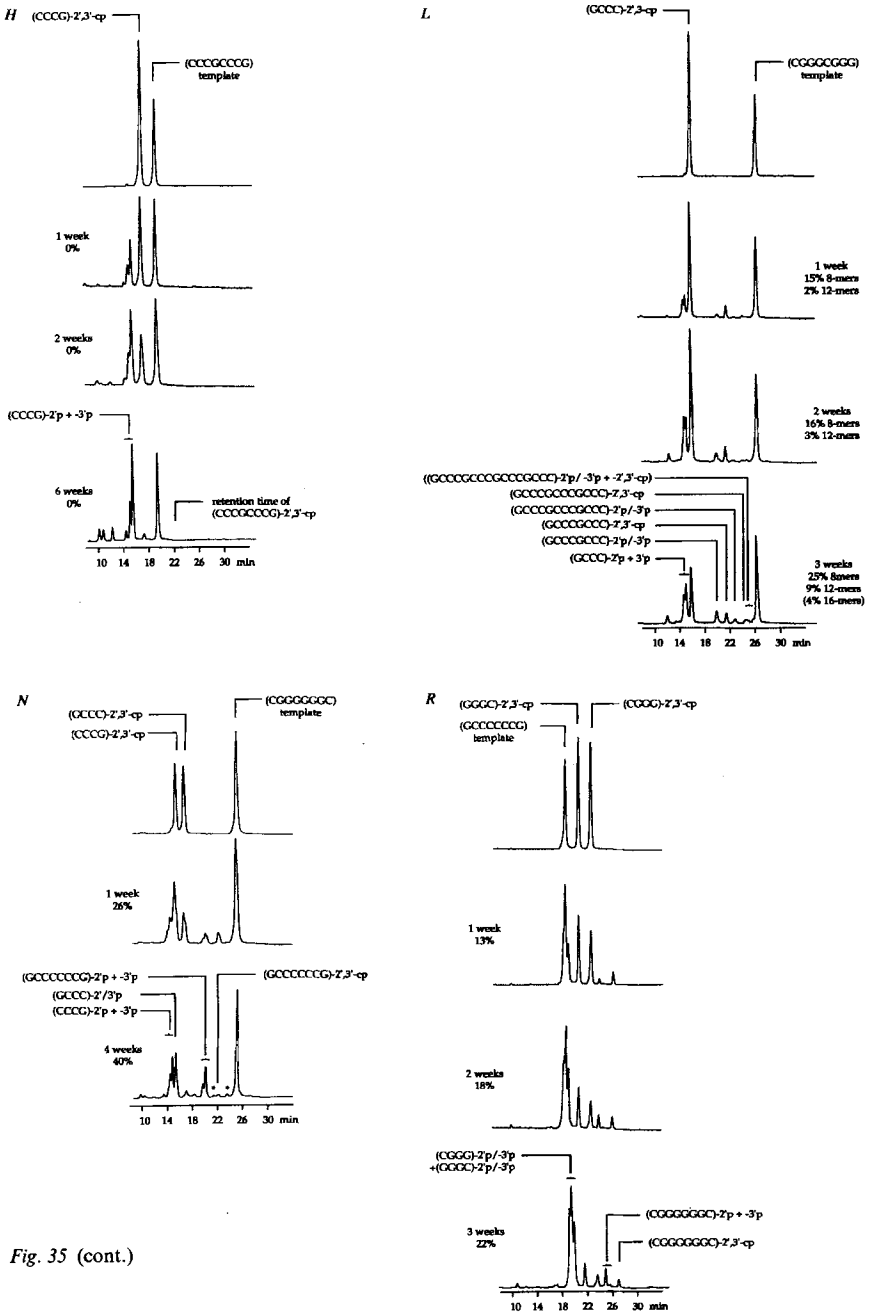

H
(CCCG)-2',3'-cp
(CCCGCCCG) template
1 week 0%
2 weeks 0%
(CCCG)-2'p + -3'p
6 weeks 0%
retention time of (CCCGCCCG)-2',3'-cp
10 14 18 22 26 30 min
L
(GCCC)-2',3-cp
(CGGGCGGG) template
1 week 15% 8-mers 2% 12-mers
2 weeks 16% 8-mers 3% 12-mers
((GCCCGCCCGCCCGCCC)-2'p/ -3'p + -2',3'-cp)
(GCCCGCCCGCCC)-2',3'-cp
(GCCCGCCCGCCC)-2'p/-3'p
(GCCCGCCC)-2',3'-cp
(GCCCGCCC)-2'p/-3'p
(GCCC)-2'p + 3'p
3 weeks 25% 8mers 9% 12-mers (4% 16-mers)
10 14 18 22 26 30 min
N
(CCCC)-2',3'-cp
(CCCG)-2',3'-cp
(CGGGGGGC) template
1 week 26%
(GCCCCCCG)-2'p + -3'p
(GCCC)-2'/3'p
(CCCG)-2'p + -3'p
(GCCCCCCG)-2',3'-cp
4 weeks 40%
10 14 18 22 26 30 min
R
(GGGC)-2',3'-cp
(GCCCCCCG) template
(CCGG)-2',3'-cp
1 week 13%
2 weeks 18%
(CGGG)-2'p/-3'p +(GCGC)-2'p/-3'p
(CGGGGGGC)-2'p + -3'p
(CGGGGGGC)-2',3'-cp
3 weeks 22%
10 14 18 22 26 30 min

Fig. 35 (cont.)

due to cyclophosphate hydrolysis. On the other hand, the amount of product phosphate monotonously increases due to hydrolysis of the ligation product cyclophosphate. To illustrate the reliability of the assignments, the HPLC traces of seven ligation experiments referred to in *Fig. 16* are given in *Fig. 35*.

Desalting of Oligonucleotides (*Sep-Pak*® cartridges). For desalting, oligonucleotide-containing aq. solns. were applied to a *Sep-Pak*® cartridge, which was washed with 10 ml of MeCN followed by 20 ml of 0.1M Et_3NHHCO_3 in H_2O prior to use. Salts were eluted from the cartridge by washing with 10–20 ml of 0.1M Et_3NHHCO_3 in H_2O. Finally, the oligonucleotide was eluted with 20–30% MeCN in H_2O. Oligonucleotide-containing fractions were combined and evaporated to dryness. An additional evaporation from 1–2 ml of H_2O removed residual Et_3NHHCO_3.

Dephosphorylation of Ligation Product 2′(and 3′)-Phosphates. Dephosphorylation was achieved by adding 0.1*n* µl of 0.1M diethanolamine buffer pH 9.8 followed by 0.01 *n* µl of alkaline phosphatase (from calf intestine, EC 3.1.3.1., 30000 U/ml, *Boehringer Mannheim, # 567744*) to *n* µl of a 2–10 µM soln. of the isolated oligonucleotide phosphate. The soln. was kept at r.t. Progress of the reaction was monitored by HPLC (*Mono Q*). In general, 2′-phosphates were completely dephosphorylated within 1–6 h, while dephosphorylation of 3′-phosphates was considerably slower (*e.g.*, D-pr(GCCCGCCC)-3′-p: 96 h). No further treatment of the dephosphorylated products was needed prior to coinjection with authentic material.

Conversion of the Ligation Product Phosphates into 2′,3′-Cyclophosphates. To a 2–10 µM soln. of the isolated and desalted ligation product phosphate fraction (in the case of (GCCCGCCC)-2′-p, the fraction also contained minor amounts of the corresponding 3′-phosphate), a large excess (150–200 mg/ml = 0.78–1.04M) of *N*-[3-(dimethylamino)propyl]-*N*′-ethylcarbodiimide hydrochloride (DEC · HCl) was added. The soln. was allowed to stand at r.t. for 1–3 h. Progress of the reaction was monitored by HPLC (*Mono Q*). For co-injections with authentic material, the reaction mixture was used without further treatment.

REFERENCES

[1] I. Schlönvogt, S. Pitsch, C. Lesueur, A. Eschenmoser, B. Jaun, R. M. Wolf, *Helv. Chim. Acta* **1996**, *79*, 2316–2345.

[2] P. Lohse, B. Oberhauser, B. Oberhauser-Hofbauer, G. Baschang, A. Eschenmoser, *Croat. Chem. Acta* **1996**, *69*, 535–562.

[3] R. Krishnamurthy, S. Pitsch, M. Minton, C. Miculka, N. Windhab, A. Eschenmoser, *Angew. Chem.* **1996**, *108*, 1619–1623; *ibid. Int. Ed. Engl.* **1996**, *35*, 1537–1541.

[4] 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*, 1621–1635.

[5] R. Micura, M. Bolli, N. Windhab, A. Eschenmoser, *Angew. Chem.* **1997**, *109*, 899–902; *ibid. Int. Ed. Engl.* **1997**, *36*, 870–873.

[6] M. Bolli, R. Micura, A. Eschenmoser, *Chem. Biol.* **1997**, *4*, 309–320.

[7] A. Eschenmoser, M. V. Kisakürek, *Helv. Chim. Acta* **1996**, *79*, 1249–1259.

[8] L. E. Orgel, *Nature (London)* **1992**, *358*, 203–209.

[9] R. Naylor, P. T. Gilham, *Biochemistry* **1966**, *5*, 2722–2728.

[10] L. E. Orgel, R. Lohrmann, *Acc. Chem. Res.* **1994**, *7*, 368–377.

[11] M. Renz, R. Lohmann, L. E. Orgel, *Biochim. Biophys. Acta* **1971**, *240*, 463–471.

[12] D. A. Usher, A. H. McHale, *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 1149–1153; D. A. Usher, *Nature (New Biol.)* **1972**, *235*, 207–208.

[13] N. G. Dolinnaya, N. I. Sokolova, D. T. Ashirbekova, Z. A. Shabarova, *Nucleic Acids Res.* **1991**, *19*, 3067–3072; N. G. Dolinnaya, I. N. Merenkova, Z. A. Shabarova, *Nucleosides Nucleotides* **1994**, *13*, 2169–2183; O. A. Federova, M. B. Gottikh, T. S. Oretskaya, Z. A. Shabarova, *ibid.* **1996**, *15*, 1137–1147.

[14] G. von Kiedrowski, *Angew. Chem. Int. Ed. Engl.* **1986**, *25*, 932–935; B. G. Bag, G. von Kiedrowski, *Pure Appl. Chem.* **1996**, *68*, 2145–2152.

[15] K. J. Luebke, P. B. Dervan, *J. Am. Chem. Soc.* **1991**, *113*, 7447–7448; K. J. Luebke, P. B. Dervan, *Nucleic Acids Res.* **1992**, *20*, 3005–3009.

[16] T. Li, K. C. Nicolaou, *Nature (London)* **1994**, *369*, 218–221.

[17] S. Pitsch, S. Wendeborn, R. Krishnamurthy, M. Bolli, A. Holzner, M. Minton, C. Micura, M. Ramaseshan, A. Eschenmoser, to be published in *Helv. Chim. Acta*.

[18] S. Pitsch, S. Wendeborn, B. Jaun, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*, 2161–2183.

[19] V. A. Efimov, A. A. Buryakova, S. V. Reverdatto, O. G. Chakhmakhcheva, Y. A. Ovchinnikov, *Nucleic Acids Res.* **1983**, *11*, 8369–8387.

[20] M. Roth, P. Dubs, E. Götschi, A. Eschenmoser, *Helv. Chim. Acta* **1971**, *54*, 710–734.
[21] A. Loupy, B. Tchoubar, 'Salt Effects in Organic and Organometallic Chemistry', VCH-Verlagsgesellschaft mbH, Weinheim, 1992, p. 73–74.
[22] D. Seebach, A. K. Beck, A. Studer, in: 'Modern Synthetic Methods', Verlag HCA, Basel/VCH-Verlagsgesellschaft mbH, Weinheim, 1995, p. 3–178.
[23] U. Pieles, W. Zürcher, M. Schär, H. E. Moser, *Nucleic Acids Res.* **1993**, *21*, 3191–3196.
[24] A. Eschenmoser, M. Dobler, *Helv. Chim. Acta* **1992**, *75*, 218–259.
[25] K. Groebke, W. Fraser, J. Hunziker, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, to be published in *Helv. Chim. Acta*.
[26] L. Tenud, S. Farooq, J. Seibl, A. Eschenmoser, *Helv. Chim. Acta* **1970**, *53*, 2059–2069.
[27] G. F. Joyce, G. M. Visser, C. A. A. van Boeckel, J. H. van Boom, L. E. Orgel, J. van Westrenen, *Nature (London)* **1984**, *310*, 602–604.
[28] M. Egli, P. Lubini, M. Bolli, S. Pitsch, A. Eschenmoser, *Chem. Biol.*, in preparation.
[29] R. Micura, M. Bolli, A. Eschenmoser, to be published.
[30] J. Hunziker, H.-J. Roth, M. Böhringer, A. Giger, U. Diederichsen, M. Göbel, Ravichandran Krishnan, B. Jaun, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*, 259–352.
[31] G. Otting, M. Billeter, K. Wüthrich, H.-J. Roth, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1993**, *76*, 2701–2756.
[32] A. Eschenmoser, 'Toward a Chemical Etioloy of the Natural Nucleic Acids' Structure', Proceed. R. A. Welch Foundation, 37th Conference on Chem. Res. '40 Years of the DNA Double Helix' 1993, p. 201–235.
[33] A. Eschenmoser, 'Warum Pentose- und nicht Hexose-Nucleinsäuren?', *Nachr. Chem. Techn. Lab.* **1991**, *39*, 795–806.
[34] A. Eschenmoser, 'Kon-Tiki-Experimente zur Frage nach dem Ursprung von Biomolekülen', in 'Materie und Prozesse – vom Elementaren zum Komplexen', Verhandlg. der Gesellschaft Deutscher Naturforscher und Ärzte, 116. Versammlung Berlin 1990, Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart, p. 135–172.
[35] M. S. Verlander, R. Lohrmann, L. E. Orgel, *J. Mol. Evol.* **1973**, *2*, 303–316.
[36] M. S. Verlander, L. E. Orgel, *J. Mol. Evol.* **1974**, *3*, 115–120.
[37] G. F. Joyce, L. E. Orgel, *J. Mol. Biol.* **1986**, *188*, 433–441.
[38] A. R. Hill, Jr., L. E. Orgel, T. Wu, *Origins Life Evol. Biosphere* **1993**, *23*, 285–290.
[39] S. Pitsch, submitted to in *Helv. Chim. Acta*.
[40] R. Kierzek, L. He, D. H. Turner, *Nucleic Acids Res.* **1992**, *20*, 1685–1690.
[41] T. Inoue, L. E. Orgel, *J. Mol. Biol.* **1982**, *162*, 201–217.
[42] T. Inoue, G. F. Joyce, K. Grzeskowiak, L. E. Orgel, J. M. Brown, C. B. Reese, *J. Mol. Biol.* **1984**, *178*, 669–676.
[43] O. L. Acevedo, L. E. Orgel, *J. Mol. Biol.* **1987**, *197*, 187–193.
[44] D. P. Bartel, J. W. Szostak, *Science* **1993**, *261*, 1411–1418.
[45] R. Rohatgi, D. P. Bartel, J. W. Szostak, *J. Am. Chem. Soc.* **1996**, *118*, 3332–3339.
[46] R. Rohatgi, D. P. Bartel, J. W. Szostak, *J. Am. Chem. Soc.* **1996**, *118*, 3340–3344.
[47] I. Tazawa, S. Tazawa, L. M. Stempel, P. O. P. Ts'o, *Biochemistry* **1970**, *9*, 3499–3514.
[48] J. P. Dougherty, C. J. Rizzo, R. Breslow, *J. Am. Chem. Soc.* **1992**, *114*, 6254–6255; R. Breslow, T. L. Sheppard, *Pure Appl. Chem.* **1996**, *68*, 2037–2041.
[49] H. Hashimoto, C. Switzer, *J. Am. Chem. Soc.* **1992**, *114*, 6255–6256; T. P. Prakash, K. E. Jung, C. Switzer, *J. Chem. Soc., Chem. Commun.* **1996**, 1793–1794.
[50] G. Ertem, J. P. Ferris, *Nature (London)* **1996**, *379*, 238–240.
[51] D. Sievers, G. von Kiedrowski, *Nature (London)* **1994**, *369*, 221–224.
[52] D. H. Lee, J. R. Granja, J. A. Martinez, K. Severin, M. R. Ghadiri, *Nature (London)* **1996**), *382*, 525–528.
[53] A. J. Hager, J. D. Pollard, Jr., J. W. Szostak, *Chem. Biol.* **1996**, *3*, 717–725; E. H. Ekland, D. P. Bartel, *Nature (London)* **1996**, *382*, 373–376.
[54] L. A. Marky, K. J. Breslauer, *Biopolymers* **1987**, *26*, 1601–1620.
[55] S. Pitsch, postdoctoral research report 1995, ETH-Zürich.
[56] N. Windhab, postdoctoral research report 1996, Institut für pharmazeutische Chemie der Universität Frankfurt/M.
[57] M. Minton, postdoctoral research report 1994, ETH-Zürich.
[58] S. Wendeborn, postdoctoral research report 1993, ETH-Zürich.