*3748–3755* *Nucleic Acids Research, 1996, Vol. 24, No. 19* 

# A convenient approach to the synthesis of trinucleotide phosphoramidites—synthons for the generation of oligonucleotide/peptide libraries

**A. L. Kayushin*, M. D. Korosteleva, A. I. Miroshnikov, W. Kosch[1], D. Zubov[1] and N. Piel[1]**

Shemiakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Mikluho-Maklaia St. 16/10, 117871 Moscow, Russia and [1]Bayer AG, Central Research, 51368 Leverkusen, Germany

Received June 18, 1996; Revised and Accepted August 20, 1996

## ABSTRACT

**Trinucleotide phosphoramidites that correspond to the codons of all 20 amino acids were synthesized in high yield in 5g scale. Precursors of those amidites—trinucleotide phosphotriesters—have been prepared using the phosphotriester approach without protection of the 3′-hydroxyl function. The structures of trinucleotide phosphotriesters and intermediates were confirmed by $^{1}$H- and $^{31}$P-NMR spectra, mass-spectra and by analysis of SPDE-hydrolysates of deprotected preparations. Purity of the target products has been confirmed by test reactions. The synthons have been used for automated synthesis of oligonucleotides and corresponding libraries by a phosphite-triester approach. A 54mer, containing 12 randomized internal bases, and a 72mer with 24 internal randomized bases have been synthesized.**

## INTRODUCTION

The key point in drug discovery programs is the screening of large numbers of different molecules for an identification of a lead compound, which possesses the desired properties. The probability of success is proportional to the size and the diversity of the screened library of compounds. Recently, large biologically displayed peptide or antibody libraries have been introduced as a source of potential pharmaceutical leads, offering a tremendous set of diverse molecules (1–3). Oligonucleotide-directed mutagenesis is the favoured method for preparing these libraries. Therefore, often subsets of the 20 amino acids are introduced at a defined position in the molecule by using oligonucleotides of a mixed composition. Such randomized oligonucleotides are usually synthesized using a mixture of commercially available nucleotides at each step of the synthesis. Despite wide use of this strategy and efforts to improve this process (4), it is nearly impossible to achieve a controlled diversity and to avoid the incorporation of undesired amino acids and stop codons in the resulting oligonucleotide. Furthermore, it is impossible to construct a desired subset of codons in a defined position. To overcome these disadvantages trinucleotide synthons may be used. Virnekas *et al*. (5) were first to use this approach.

Here we describe a modified method for the synthesis of such trinucleotide phosphoramidites building blocks. We have succeeded to scale up the reactions and to use the products in a conventional automated solid-phase synthesis of nucleotides. Two examples are listed illustrating the advantages and disadvantages of the library syntheses. Furthermore, we have built up codons different to those of Virnekas *et al*. In our view, they are more suitable to be used for library construction in phage or *Escherichia coli* systems.

## EXPERIMENTAL

### General materials and methods

Base-protected dimethoxytrityl nucleosides were obtained from ‘Vostok Ltd’ (Russia). 3-Nitro-1,2,4-triazole and benzenesulphonic acid (Fluka) were used without purification. *o*-Chlorophenol, $POCl_3$, $PCl_3$, diisopropylamine and cyanoethanol (Fluka) were distilled before use. Triazole (Fluka) was recrystallized from toluene. Mesitylenesulphonylchloride (Fluka) was recrystallized from pentane. Tetrazole (Fluka) was sublimed *in vacuo* at 120°C. Acetonitrile and pyridine were distilled over calcium hydride and stored over molecular sieves (4A, Serva). 1-Mesitylenesulphonyl-3-nitro-triazole (MSNT) was prepared as previously described (6; recrystallized from dry benzene, yield 80%, m.p. 138–140°C). The phosphitylating reagent **7** was prepared according to Nielsen *et al*. (7). Snake venom phosphodiesterase (SPDE, EC 3.1.4.1), $MgCl_2$, dithiothreitol and Tris were purchased from Sigma. TLC was performed on silica-gel plates (Kieselgel 60 F254, Merck) using $CHCl_3$–MeOH (9:1) as a developing system. Silica gel 60 (Fluka) and Separon SGX C18 60 (Ltd Praha, Czechoslovakia) were used for column chromatography. The HPLC was performed on a Gilson system (models 305 and 303) using the Kratos UV-detector Spectroflow 757. Zorbax-ODS column 4.6 × 250 mm (DuPont Instruments) was used for HPLC-analysis of preparations (gradient 75–100% acetonitrile in water, 60 min, flow rate 1 ml/min, detection at 280 nm). The HPLC analysis of unprotected di- and trinucleotides and of SPDE-hydrolysates was performed on Nucleosil C18 column 4.6 × 250 mm (Macherey-Nagel) (0–50% acetonitrile in 0.1 M NaOAc, pH 6.6, 60 min, flow rate 1 ml/min, 270 nm). The

* To whom correspondence should be addressed

oligonucleotides were synthesized using a DNA-synthesizer 380B (Applied Biosystems).

$^{1}$H-NMR spectra were measured at 600 MHz using a Varian UNITY 600 spectrometer. $^{31}$P-NMR spectra were measured at 202.48 MHz using a Bruker WM-500 spectrometer. Chemical shifts are given in p.p.m. relative to tetramethylsilane and 85% $H_3PO_4$ correspondingly as an external standard.

The time-of-flight $^{252}$Cf-plasma desorption mass-spectrometer MSBC ('Electron Ltd', Ukraine) was used for measurements of mass-spectra.

## Syntheses

*Synthesis of o-chlorophenylphosphodichloridate. o*-Chlorophenol (74 ml, 0.7 mol) was added dropwise within 1 h under intensive stirring at 90–100°C to a solution of 700 mg anhydrous $AlCl_3$ in 78 ml (1.4 mol) $POCl_3$. The resulting solution was heated at 110°C for 1 h and cooled to room temperature. Excess $POCl_3$ was removed by distillation under reduced pressure and the residue was distilled twice under high vacuum, yielding 75 g (0.35 mol, 45%) of the target product.

*Synthesis of 5′-O-dimethoxytrityl-N-acyl-2′-deoxynucleoside-3′-O-(2-chlorophenylphosphate) (triethylammonium salt)* (**1**). Triazole (7 g, 102 mmol) and anhydrous triethylamine (8 ml, 101 mmol) were dissolved in 160 ml of dioxane and cooled to 5°C. A solution of 12.4 g (50 mmol) of *o*-chlorophenylphosphodichloridate in 25 ml dioxane was added dropwise. After 1 h the solution was filtered into the cooled to nearly –5°C solution of 25 mmol 5′-*O*-dimethoxytrityl-*N*-acyl-2′-deoxynucleoside (dried by a coevaporation with dry pyridine—3 × 100 ml) in 25 ml dry pyridine. After completion of the reaction (1 h) 250 ml 0.1 M TEAB was added under cooling. The solution was concentrated, the residue was dissolved in 100 ml $CHCl_3$ and washed by 100 ml 0.1 M TEAB. The aqueous layer was washed with 100 ml of chloroform. The combined organic layer was washed by 0.1 M TEAB (2 × 200 ml), dried ($Na_2SO_4$), and concentrated to nearly 50 ml. Compound **1** was precipitated with 2.5 l pentane in a yield of 90–95%.

*Synthesis of 5′-O-dimethoxytritylated di- and trinucleotides* (**3a** and **5**). General procedure: the mixture of P- and OH-components was coevaporated with dry pyridine (3 × 50 ml). A freshly prepared solution of MSNT in dry pyridine was added and the reaction mixture was concentrated to the reaction volume (nearly 2 ml pyridine per 1 mmol of each component, including MSNT). After completion (1 h, room temperature, monitored by TLC) the reaction mixture was cooled to 0°C and an equal volume of water was added. After 15 min the mixture was concentrated to a gum, dissolved in $CHCl_3$ (nearly 200 ml) and washed with 0.1 M TEAB. The aqueous layer was washed with 100 ml of $CHCl_3$. The combined organic layers were washed with 0.1 M TEAB (3 × 300 ml), dried ($Na_2SO_4$) and concentrated. The pyridine was removed by coevaporation with toluene (3 × 50 ml), the residual gum was dissolved in $CHCl_3$ and concentrated to a minimal volume. The product was isolated by column chromatography on silica gel. For scale up to 10 mmol of OH-component the column 4 × 30 cm was used, elution by gradient of methanol in $CHCl_3$ (0–12%, 2 l, flow rate 1 l/h). For larger reaction scales the column 6 × 40 cm was used (elution by the same system, 5 l, flow rate 2 l/h). Component ratios and reaction yields are listed in Tables 2 and 3.

$^{1}$H-NMR. For all dinucleotides **3a**: δ 2.24–2.41 (4H, m, 2′-H), 3.74 (6H, s, $CH_3O$ of DTr), 3.75–3.77 (2H, m, 5′-$CH_2$ of 5′-nucleoside), 4.44–4.50 (2H, m, 4′-H), 4.52–4.61 (2H, m, 5′-$CH_2$ of 3′-nucleoside), 5.25–5.45 (2H, m, 3′-H), 6.15–6.24 (2H, m, 1′-H), 6.70–6.75 (4H, m, 3,3',5,5′-H of DTr), 7.13–7.23 (13H, m, 3,4,5,6-H of ClPh, 2,2′,6,6′,2″,3″,4″,5″,6″-H of DTr).

AA: δ 7.36–7.54 (6H, m, 3,4,5-H of Bz), 7.92–8.02 (4H, m, 2,6-H of Bz), 8.04–8.22 (2H, m, 2-H of Ade), 8.58–8.90 (2H, m, 8-H of Ade), 10.43–10.90 (2H, m, NH).

AC: δ 7.04–7.12 (2H, m, 5,6-H of Cyt), 7.35–7.56 (6H, m, 3,4,5-H of Bz), 7.93–8.03 (4H, m, 2,6-H of Bz), 8.04–8.22 (1H, m, 2-H of Ade), 8.60–8.90 (1H, m, 8-H of Ade), 10.42–10.87 (2H, m, NH).

AG: δ 1.01–1.12 (6H, m, $CH_3$ of i-Bu), 2.28–2.34 (1H, m, CH of i-Bu), 7.38–7.52 (3H, m, 3,4,5-H of Bz), 7.72 (1H, s, 8-H of Gua), 7.77–8.04 (2H, m, 2,6-H of Bz), 8.16–8.27 (1H, m, 2-H of Ade), 8.73–8.80 (1H, m, 8-H of Ade), 10.27–10.66 (2H, m, NH).

AT: δ 1.68–1.76 (3H, m, $CH_3$ of Thy), 7.04–7.15 (1H, m, 6-H of Thy), 7.34–7.60 (3H, m, 3,4,5-H of Bz), 7.96–8.06 (2H, m, 2,6-H of Bz), 8.06–8.22 (1H, m, 2-H of Ade), 8.63–8.73 (1H, m, 8-H of Ade), 10.27–10.94 (1H, m, NH).

CG: δ 0.97–1.11 (6H, m, $CH_3$ of i-Bu), 2.19–2.24 (1H, m, CH of i-Bu), 7.04–7.13 (2H, m, 5,6-H of Cyt), 7.32–7.50 (3H, m, 3,4,5-H of Bz), 7.71–7.74 (1H, m, 8-H of Gua), 7.81–7.94 (2H, m, 2,6-H of Bz), 10.24–10.71 (2H, m, NH).

CT: δ 1.67–1.74 (3H, m, $CH_3$ of Thy), 7.07–7.13 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.33–7.45 (3H, m, 3,4,5-H of Bz), 7.91–8.04 (2H, m, 2,6-H of Bz), 10.35–10.97 (1H, m, NH).

GC: δ 0.95–1.07 (6H, m, $CH_3$ of i-Bu), 2.29–2.48 (1H, m, CH of i-Bu), 7.07–7.16 (2H, m, 5,6-H of Cyt), 7.40–7.50 (3H, m, 3,4,5-H of Bz), 7.75–7.77 (1H, m, 8-H of Gua), 7.79–7.86 (2H, m, 2,6-H of Bz), 10.39–10.70 (2H, m, NH).

GG: δ 1.07–1.11 (12H, m, $CH_3$ of i-Bu), 2.32–2.49 (2H, m, CH of i-Bu), 7.76 (2H, s, 8-H of Gua), 10.19–10.56 (2H, m, NH).

GT: δ 0.97–1.12 (6H, m, $CH_3$ of i-Bu), 1.70–1.77 (3H, m, $CH_3$ of Thy), 2.22–2.44 (1H, m, CH of i-Bu), 7.07–7.13 (1H, m, 6-H of Thy), 7.70–7.72 (1H, m, 8-H of Gua), 10.19–10.37 (1H, m, NH).

TC: δ 1.76–1.82 (3H, m, $CH_3$ of Thy), 7.07–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.34–7.61 (3H, m, 3,4,5-H of Bz), 7.95–8.05 (2H, m, 2,6-H of Bz), 10.42–11.02 (1H, m, NH).

TG: δ 0.99–1.12 (6H, m, $CH_3$ of i-Bu), 1.66–1.73 (3H, m, $CH_3$ of Thy), 2.20–2.41 (1H, m, CH of i-Bu), 7.06–7.14 (1H, m, 6-H of Thy), 7.72 (1H, s, 8-H of Gua), 10.11–10.25 (1H, m, NH).

TT: δ 1.71–1.78 (6H, m, $CH_3$ of Thy), 7.07–7.13 (2H, m, 6-H of Thy).

For all trinucleotides **5**: δ 2.22–2.42 (6H, m, 2′-H), 3.74 (6H, s, $CH_3O$ of DTr), 3.75–3.78 (2H, m, 5′-$CH_2$ of 5′-nucleoside), 4.41–4.47 (3H, m, 4′-H), 4.47–4.61 (4H, m, internucleotide 5′-$CH_2$), 5.17–5.46 (3H, m, 3′-H), 6.14–6.31 (3H, m, 1′-H), 6.76–6.81 (4H, m, 3,3′,5,5′-H of DTr), 7.14–7.26 (17H, m, 3,4,5,6-H of ClPh, 2,2′,6,6′,2″,3″,4″,5″,6″-H of DTr).

AAA: δ 7.36–7.57 (9H, m, 3,4,5-H of Bz), 7.94–8.08 (6H, m, 2,6-H of Bz), 8.08–8.21 (3H, m, 2-H of Ade), 8.60–8.94 (3H, m, 8-H of Ade), 10.43–10.87 (3H, m, NH).

AAC: δ 7.05–7.10 (2H, m, 5,6-H of Cyt), 7.34–7.55 (9H, m, 3,4,5-H of Bz), 7.93–8.00 (6H, m, 2,6-H of Bz), 8.05–8.22 (2H, m, 2-H of Ade), 8.56–8.85 (2H, m, 8-H of Ade), 10.50–11.00 (3H, m, NH).

ACT: δ 1.70–1.80 (3H, m, $CH_3$ of Thy), 7.07–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.32–7.46 (6H, m, 3,4,5-H of Bz), 7.91–8.05 (4H, m, 2,6-H of Bz), 8.13–8.25 (1H, m, 2-H of Ade), 8.62–8.69 (1H, m, 8-H of Ade), 10.35–11.00 (2H, m, NH).

ATC: δ 1.75–1.83 (3H, m, $CH_3$ of Thy), 7.06–7.13 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.33–7.61 (6H, m, 3,4,5-H of Bz), 7.95–8.07 (4H, m, 2,6-H of Bz), 8.09–8.22 (1H, m, 2-H of Ade), 8.62–8.69 (1H, m, 8-H of Ade), 10.44–11.02 (2H, m, NH).

ATG: δ 1.04–1.18 (6H, m, $CH_3$ of i-Bu), 1.64–1.77 (3H, m, $CH_3$ of Thy), 2.23–2.34 (1H, m, $CH_3$ of i-Bu), 7.07–7.15 (1H, m, 6-H of Thy), 7.34–7.56 (3H, m, 3,4,5-H of Bz), 7.67–7.71 (1H, m, 8-H of Gua), 7.98–8.08 (2H, m, 2,6-H of Bz), 8.09–8.26 (1H, m, 2-H of Ade), 8.63–8.72 (1H, m, 8-H of Ade), 10.53–10.95 (2H, m, NH).

CAG: δ 1.01–1.14 (6H, m, $CH_3$ of i-Bu), 2.27–2.35 (1H, m, CH of i-Bu), 7.10–7.15 (1H, m, 6-H of Cyt), 7.39–7.55 (6H, m, 3,4,5-H of Bz), 7.66–7.69 (1H, m, 8-H of Gua), 7.78–8.00 (4H, m, 2,6-H of Bz), 8.16–8.28 (1H, m, 2-H of Ade), 8.73–8.87 (1H, m, 8-H of Ade), 10.42–10.90 (3H, m, NH).

CAT: δ 1.61–1.68 (3H, m, $CH_3$ of Thy), 7.06–7.13 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.32–7.62 (6H, m, 3,4,5-H of Bz), 7.97–8.03 (4H, m, 2,6-H of Bz), 8.05–8.23 (1H, m, 2-H of Ade), 8.63–8.72 (1H, m, 8-H of Ade), 10.16–10.87 (2H, m, NH).

CCG: δ 0.96–1.11 (6H, m, $CH_3$ of i-Bu), 2.19–2.22 (1H, m, CH of i-Bu), 7.06–7.12 (4H, m, 5,6-H of Cyt), 7.34–7.53 (6H, m, 3,4,5-H of Bz), 7.71–7.74 (1H, m, 8-H of Gua), 7.83–7.97 (4H, m, 2,6-H of Bz), 10.11–10.23 (3H, m, NH).

CGT: δ 0.98–1.12 (6H, m, $CH_3$ of i-Bu), 1.69–1.75 (3H, m, $CH_3$ of Thy), 2.20–2.41 (1H, m, CH of i-Bu), 7.06–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.33–7.54 (3H, m, 3,4,5-H of Bz), 7.69–7.73 (1H, m, 8-H of Gua), 7.83–8.01 (2H, m, 2,6-H of Bz), 10.22–10.36 (2H, m, NH).

CTG: δ 0.99–1.13 (6H, m, $CH_3$ of i-Bu), 1.66–1.70 (3H, m, $CH_3$ of Thy), 2.22–2.42 (1H, m, CH of i-Bu), 7.06–7.16 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.33–7.51 (3H, m, 3,4,5-H of Bz), 7.72 (1H, s, 8-H of Gua), 7.84–7.92 (2H, m, 2,6-H of Bz), 10.11–10.26 (2H, m, NH).

GAA: δ 1.00–1.08 (6H, m, $CH_3$ of i-Bu), 2.21–2.41 (1H, m, CH of i-Bu), 7.31–7.61 (6H, m, 3,4,5-H of Bz), 7.67–7.72 (1H, m, 8-H of Gua), 7.83–7.95 (4H, m, 2,6-H of Bz), 8.08–8.26 (2H, m, 2-H of Ade), 8.70–8.79 (2H, m, 8-H of Ade), 10.21–10.63 (3H, m, NH).

GAC: δ 1.00–1.13 (6H, m, $CH_3$ of i-Bu), 2.38–2.48 (1H, m, CH of i-Bu), 7.11–7.16 (1H, m, 6-H of Cyt), 7.39–7.53 (6H, m, 3,4,5-H of Bz), 7.68–7.70 (1H, m, 8-H of Gua), 7.76–7.99 (4H, m, 2,6-H of Bz), 8.16–8.26 (1H, m, 2-H of Ade), 8.73–8.86 (1H, m, 8-H of Ade), 10.41–10.89 (3H, m, NH).

GCT: δ 0.93–1.09 (6H, m, $CH_3$ of i-Bu), 1.68–1.72 (3H, m, $CH_3$ of Thy), 2.19–2.42 (1H, m, CH of i-Bu), 7.08–7.16 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.39–7.46 (3H, m, 3,4,5-H of Bz), 7.64–7.70 (1H, m, 8-H of Gua), 7.66–7.83 (2H, m, 2,6-H of Bz), 10.33–10.73 (2H, m, NH).

GGT: δ 1.03–1.07 (12H, m, $CH_3$ of i-Bu), 1.74–1.77 (3H, m, $CH_3$ of Thy), 2.30–2.42 (2H, m, CH of i-Bu), 7.03–7.12 (1H, m, 6-H of Thy), 7.75–7.77 (2H, m, 8-H of Gua), 10.43–10.83 (2H, m, NH).

GTT: δ 0.94–1.11 (6H, m, $CH_3$ of i-Bu), 1.71–1.80 (6H, m, $CH_3$ of Thy), 2.14–2.38 (1H, m, CH of i-Bu), 7.08–7.15 (2H, m, 6-H of Thy), 7.64–7.68 (1H, m, 8-H of Gua), 10.41–10.70 (1H, m, NH).

TAC: δ 1.77–1.84 (3H, m, $CH_3$ of Thy), 7.02–7.10 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.31–7.49 (6H, m, 3,4,5-H of Bz), 7.89–8.05 (4H, m, 2,6-H of Bz), 8.15–8.27 (1H, m, 2-H of Ade), 8.66–8.76 (1H, m, 8-H of Ade), 10.40–10.98 (2H, m, NH).

TCT: δ 1.65–1.73 (6H, m, $CH_3$ of Thy), 7.06–7.15 (4H, m, 5,6-H of Cyt, 6-H of Thy), 7.42–7.52 (3H, m, 3,4,5-H of Bz), 7.79–7.88 (2H, m, 2,6-H of Bz), 10.52–11.00 (1H, m, NH).

TGC: δ 0.94–1.07 (6H, m, $CH_3$ of i-Bu), 1.85–1.88 (3H, m, $CH_3$ of Thy), 2.30–2.50 (1H, m, CH of i-Bu), 7.06–7.17 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.41–7.52 (3H, m, 3,4,5-H of Bz), 7.74–7.77 (1H, m, 8-H of Gua), 7.78–7.87 (2H, m, 2,6-H of Bz), 10.42–10.71 (2H, m, NH).

TGG: δ 1.05–1.10 (12H, m, $CH_3$ of i-Bu), 1.80–1.88 (3H, m, $CH_3$ of Thy), 2.30–2.48 (2H, m, CH of i-Bu), 7.01–7.14 (1H, m, 6-H of Thy), 7.75–7.78 (2H, m, 8-H of Gua), 10.21–10.59 (2H, m, NH).

TTC: δ 1.58–1.68 (6H, m, $CH_3$ of Thy), 7.06–7.15 (4H, m, 5,6-H of Cyt, 6-H of Thy), 7.42–7.50 (3H, m, 3,4,5-H of Bz), 7.80–7.88 (2H, m, 2,6-H of Bz), 10.30–10.72 (1H, m, NH).

*Removal of the dimethoxytrityl group*. A 5′-*O*-dimethoxytritylated dinucleotide **3a** (5 mmol) was dissolved in 90 ml of a 2% solution of benzenesulphonic acid in $CHCl_3$/methanol (7:3 v/v) at 0°C. After 3 min 120 ml of 5% aqueous $NaHCO_3$ was added. At this step an oily suspension was obtained. The aqueous layer was washed by 10% methanol in $CHCl_3$ until clearness. The organic layer was evaporated and the detritylated **3a** was isolated in 95–98% yield as described for the starting material, except of the elution gradient (0–15% MeOH in $CHCl_3$).

*Enzymatic analysis of di- and trinucleotides* **3** *and* **5**. Di- or trinucleotide **3** or **5** (1 mg) was dissolved in 500 µl of 30% aqueous $NH_3$ and heated at 50°C overnight. The solution was evaporated to dryness, then 100 µl of 80% acetic acid and, after 20 min, 100 µl of 10 M Tris were added, followed by HPLC. After lyophilization the product was dissolved in 500 µl water. Nearly 0.1 $AU_{260}$ of di- or trinucleotide was incubated (37°C, 3 h) with snake venom phosphodiesterase (100 µg/ml) in 100 µl buffer (10 mM $MgCl_2$, 1 mM dithiothreitol, 60 mM Tris–HCl, pH 7.5). The reaction mixture was analyzed by HPLC as described for the starting material. A solution of nucleoside and 5′ nucleotide(s) mixed in the expected ratio and incubated with phosphodiesterase under the same conditions was used as a reference. The results of the analysis are shown in Tables 2 and 3.

*Synthesis of trinucleotide phosphoramidites* (**6**). The solution of trinucleotide **5** and tetrazole in pyridine was concentrated and the residue was coevaporated with dry acetonitrile (nearly 100 ml) up to the reaction volume (50–60 ml). If the solution stayed clear, it was concentrated and coevaporated with acetonitrile again; the reaction flask was filled with argon after each coevaporation. As a result, a slightly dimmed solution of trinucleotide **5** and tetrazole in the 50–60 ml mixture of pyridine and acetonitrile was obtained. The ratio of pyridine and acetonitrile (and the corresponding number of coevaporations) strongly depends on the sequence of the trinucleotide and has not been analyzed. Phosphitylating reagent **7** was added dropwise to the solution under intensive shaking. After completion of the reaction (45 min, monitored by HPLC) the mixture was concentrated and the residue was dissolved in 200 ml of $CHCl_3$, saturated with the saturated aqueous $NaHCO_3$. The organic layer was washed with saturated $NaHCO_3$ (3 × 200 ml), dried ($Na_2SO_4$), concentrated and coevaporated with acetonitrile (2 × 100 ml). Water was added to the residual gum up to dimness and the resulting solution was applied to a column containing Separon SGX C18 (3 × 33 cm), equilibrated by 75% acetonitrile in water. First, the column was developed by 200 ml 75% acetonitrile (flow rate 390 ml/h), then by a gradient of acetonitrile in water (75–100%, 1 l) followed by the pure acetonitrile. The solution containing the target product was concentrated *in vacuo*, coevaporated with acetonitrile (5 × 200 ml) and dried under high vacuum. Component ratios, reaction yields and description of $^{31}$P-NMR spectra are presented in Table 4.

$^1$H-NMR. For all trinucleotides **6**: δ 1.14–1.19 (12H, 2d or m, $CH_3$ of i-Pr), 2.25–2.43 (6H, m, 2′-H), 2.60–2.64 (2H, t or m, $CH_2CN$), 2.73–2.85 (2H, 2t or m, $POCH_2$), 3.37–3.48 (2H, m, CH of i-Pr), 3.74 (6H, s, $CH_3O$ of DTr), 3.75–3.78 (2H, m, 5′-$CH_2$ of 5′-nucleoside), 4.43–4.48 (3H, m, 4′-H), 4.50–4.60 (4H, m, internucleotide 5′-$CH_2$), 5.27–5.46 (3H, m, 3′-H), 6.15–6.26 (3H, m, 1′-H), 6.71–6.75 (4H, m, 3,3′,5,5′-H of DTr), 7.14–7.24 (17H, m, 3,4,5,6-H of ClPh, 2,2′,6,6′,2″,3″,4″,5″, 6″-H of DTr).

AAA: δ 7.35–7.56 (9H, m, 3,4,5-H of Bz), 7.94–8.06 (6H, m, 2,6-H of Bz), 8.08–8.20 (3H, m, 2-H of Ade), 8.57–8.79 (3H, m, 8-H of Ade), 10.53–10.98 (3H, m, NH).

AAC: δ 7.09–7.13 (2H, m, 5,6-H of Cyt), 7.34–7.57 (9H, m, 3,4,5-H of Bz), 7.94–8.05 (6H, m, 2,6-H of Bz), 8.06–8.22 (2H, m, 2-H of Ade), 8.56–8.85 (2H, m, 8-H of Ade), 10.53–10.98 (3H, m, NH).

ACT: δ 1.76–1.84 (3H, m, $CH_3$ of Thy), 7.09–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.32–7.46 (6H, m, 3,4,5-H of Bz), 7.92–8.04 (4H, m, 2,6-H of Bz), 8.14–8.24 (1H, m, 2-H of Ade), 8.61–8.67 (1H, m, 8-H of Ade), 10.46–11.12 (2H, m, NH).

ATC: δ 1.73–1.82 (3H, m, $CH_3$ of Thy), 7.08–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.35–7.58 (6H, m, 3,4,5-H of Bz), 7.96–8.06 (4H, m, 2,6-H of Bz), 8.09–8.22 (1H, m, 2-H of Ade), 8.61–8.67 (1H, m, 8-H of Ade), 10.46–11.12 (2H, m, NH).

ATG: δ 1.04–1.17 (6H, m, $CH_3$ of i-Bu), 1.72–1.85 (3H, m, $CH_3$ of Thy), 2.23–2.32 (1H, m, CH of i-Bu), 7.10–7.16 (1H, m, 6-H of Thy), 7.35–7.55 (3H, m, 3,4,5-H of Bz), 7.66–7.71 (1H, m, 8-H of Gua), 7.97–8.04 (2H, m, 2,6-H of Bz), 8.08–8.25 (1H, m, 2-H of Ade), 8.64–8.74 (1H, m, 8-H of Ade), 10.53–10.93 (2H, m, NH).

CAG: δ 1.01–1.13 (6H, m, $CH_3$ of i-Bu), 2.37–2.45 (1H, m, CH of i-Bu), 7.12–7.17 (1H, m, 6-H of Cyt), 7.39–7.53 (6H, m, 3,4,5-H of Bz), 7.64–7.66 (1H, m, 8-H of Gua), 7.78–7.96 (4H, m, 2,6-H of Bz), 8.18–8.27 (1H, m, 2-H of Ade), 8.71–8.82 (1H, m, 8-H of Ade), 10.53–11.02 (3H, m, NH).

CAT: δ 1.72–1.81 (3H, m, $CH_3$ of Thy), 7.10–7.15 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.33–7.59 (6H, m, 3,4,5-H of Bz), 7.98–8.04 (4H, m, 2,6-H of Bz), 8.07–8.22 (1H, m, 2-H of Ade), 8.61–8.68 (1H, m, 8-H of Ade), 10.45–11.16 (2H, m, NH).

CCG: δ 0.96–1.10 (6H, m, $CH_3$ of i-Bu), 2.21–2.33 (1H, m, CH of i-Bu), 7.06–7.12 (4H, m, 5,6-H of Cyt), 7.34–7.52 (6H, m, 3,4,5-H of Bz), 7.68–7.72 (1H, m, 8-H of Gua), 7.84–7.96 (4H, m, 2,6-H of Bz), 10.40–10.51 (3H, m, NH).

CGT: δ 0.98–1.11 (6H, m, $CH_3$ of i-Bu), 1.78–1.85 (3H, m, $CH_3$ of Thy), 2.22–2.32 (1H, m, CH of i-Bu), 7.06–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.35–7.54 (3H, m, 3,4,5-H of Bz), 7.66–7.71 (1H, m, 8-H of Gua), 7.84–7.99 (2H, m, 2,6-H of Bz), 10.42–10.56 (2H, m, NH).

CTG: δ 0.99–1.12 (6H, m, $CH_3$ of i-Bu), 1.77–1.84 (3H, m, $CH_3$ of Thy), 2.24–2.32 (1H, m, CH of i-Bu), 7.06–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.33–7.53 (3H, m, 3,4,5-H of Bz), 7.67–7.71 (1H, m, 8-H of Gua), 7.85–8.00 (2H, m, 2,6-H of Bz), 10.40–10.55 (2H, m, NH).

GAA: δ 1.00–1.07 (6H, m, $CH_3$ of i-Bu), 2.23–2.31 (1H, m, CH of i-Bu), 7.31–7.62 (6H, m, 3,4,5-H of Bz), 7.67–7.72 (1H, m, 8-H of Gua), 7.84–7.97 (4H, m, 2,6-H of Bz), 8.09–8.28 (2H, m, 2-H of Ade) 8.71–8.78 (2H, m, 8-H of Ade), 10.50–10.92 (3H, m, NH).

GAC: δ 1.00–1.13 (6H, m, $CH_3$ of i-Bu), 2.38–2.46 (1H, m, CH of i-Bu), 7.11–7.17 (1H, m, 6-H of Cyt), 7.39–7.52 (6H, m, 3,4,5-H of Bz), 7.66–7.68 (1H, m, 8-H of Gua), 7.77–7.97 (4H, m, 2,6-H of Bz), 8.18–8.27 (1H, m, 2-H of Ade), 8.71–8.81 (1H, m, 8-H of Ade), 10.53–11.00 (3H, m, NH).

GCT: δ 0.93–1.08 (6H, m, $CH_3$ of i-Bu), 1.79–1.86 (3H, m, $CH_3$ of Thy), 2.21–2.32 (1H, m, CH of i-Bu), 7.10–7.14 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.39–7.41 (3H, m, 3,4,5-H of Bz), 7.61–7.67 (1H, m, 8-H of Gua), 7.67–7.79 (2H, m, 2,6-H of Bz), 10.53–10.93 (2H, m, NH).

GGT: δ 1.03–1.06 (12H, m, $CH_3$ of i-Bu), 1.85–1.91 (3H, m, $CH_3$ of Thy), 2.32–2.41 (2H, m, CH of i-Bu), 7.03–7.11 (1H, m, 6-H of Thy), 7.73–7.75 (2H, m, 8-H of Gua), 10.53–10.93 (2H, m, NH).

GTT: δ 0.95–1.10 (6H, m, $CH_3$ of i-Bu), 1.78–1.85 (6H, m, $CH_3$ of Thy), 2.16–2.30 (1H, m, CH of i-Bu), 7.10–7.13 (2H, m, 6-H of Thy), 7.62–7.66 (1H, m, 8-H of Gua), 10.43–10.75 (1H, m, NH).

TAC: δ 1.77–1.84 (3H, m, $CH_3$ of Thy), 7.08–7.13 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.30–7.44 (6H, m, 3,4,5-H of Bz), 7.89–8.00 (4H, m, 2,6-H of Bz), 8.18–8.26 (1H, m, 2-H of Ade), 8.64–8.71 (1H, m, 8-H of Ade), 10.50–11.16 (2H, m, NH).

TCT: δ 1.73–1.82 (6H, m, $CH_3$ of Thy), 7.06–7.16 (4H, m, 5,6-H of Cyt, 6-H of Thy), 7.42–7.51 (3H, m, 3,4,5-H of Bz), 7.80–7.90 (2H, m, 2,6-H of Bz), 10.52–10.94 (1H, m, NH).

TGC: δ 0.94–1.06 (6H, m, $CH_3$ of i-Bu), 1.95–1.99 (3H, m, $CH_3$ of Thy), 2.32–2.42 (1H, m, CH of i-Bu), 7.07–7.18 (3H, m, 5,6-H of Cyt, 6-H of Thy), 7.41–7.50 (3H, m, 3,4,5-H of Bz), 7.71–7.74 (1H, m, 8-H of Gua), 7.79–7.86 (2H, m, 2,6-H of Bz), 10.42–10.81 (2H, m, NH).

TGG: δ 1.04–1.08 (12H, m, $CH_3$ of i-Bu), 1.86–1.92 (3H, m, $CH_3$ of Thy), 2.30–2.41 (2H, m, CH of i-Bu), 7.03–7.12 (1H, m, 6-H of Thy), 7.72–7.76 (2H, m, 8-H of Gua), 10.50–10.88 (2H, m, NH).

TTC: δ 1.68–1.80 (6H, m, $CH_3$ of Thy), 7.06–7.15 (4H, m, 5,6-H of Cyt, 6-H of Thy), 7.42–7.52 (3H, m, 3,4,5-H of Bz), 7.81–7.90 (2H, m, 2,6-H of Bz), 10.40–10.81 (1H, m, NH).

*Analysis of phosphoramidites* **6** *reactivity.* A 10% solution of phosphoramidite in acetonitrile and 10% solution of tetrazole in methanol (50 µl of each) were mixed. After 45 s the reaction mixture was analyzed under the same conditions as the starting phosphoramidite **6**.

*Synthesis of DNA-libraries.* DNA libraries have been synthesized by a traditional phosphoramidite approach. A slightly modified synthesis cycle 'CEAF-3' of the manufacturer was used. The following changes were made: double coupling of the monomers; triple coupling of the trimers; extension of the coupling time of the trimers to 120 s; increase of amount of trimers added to the coupling mixture (+2 s). The trimers were dissolved in an acetonitrile/dichloromethane mixture (1:3 v/v). The resulting concentration of trimers was 0.15 M while their ratios were defined by the 'reaction factors' (see Results and Discussion).

The DNAs were cleaved from the support by 25% aqueous ammonia and the resulted solution was kept at room temperature overnight and then heated to 55°C for additional 4 h followed by evaporation of the solvent. The 5′-*O*-dimethoxytritylated DNA library was purified by HPLC (elution solvents: A, 0.1 M triethylammoniumacetate in water, pH 7.0; B, acetonitrile; 5% B to 60% B, 30 min; column: Hypersil RP 18, 8×250 mm, Knauer, Germany). After detritylation (80% acetic acid in water, 5 min at 55°C and 25 min at room temperature) the desired products were isolated by HPLC under the same conditions.

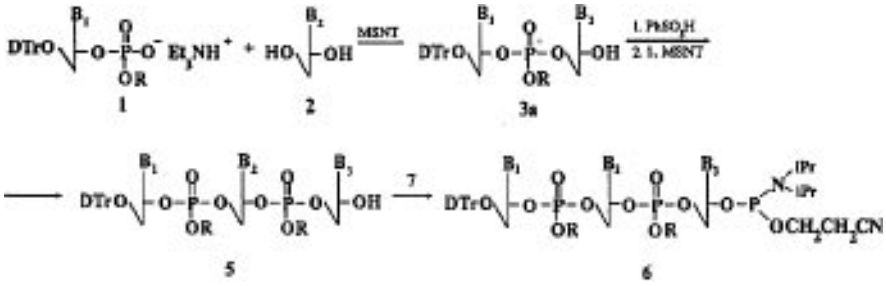

**Figure 1.** The synthetic route of trinucleotide phosphoramidites.

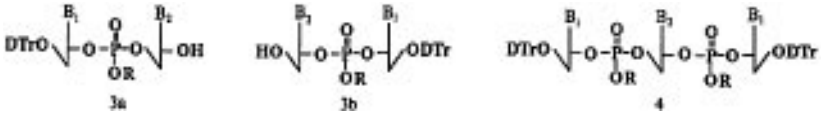

**Figure 2.** The possible by-products after synthesis of 3′-unprotected dinucleotide.

## RESULTS AND DISCUSSION

The general plan of the trinucleotide synthesis consisted of two main parts. First, we had to decide which trinucleotides to synthesize.

We planned to use *E.coli* as the host organism. This system is widely used for phage display construction (10,11), recombinant antibody production (12) and surface display libraries (13). The level of expression of recombinant proteins containing randomized fragments in all those systems is generally very high. Clearly, rarely used codons in randomized fragments must be omitted to avoid restrictions on the translation level.

Our choice of trinucleotide structures was based on codon usage in *E.coli* (8,9) rather than on simplification of the synthesis strategy. Hence there is the substantial difference between our base sequences and those described by Virnekas *et al.* (5). The structure of codons chosen and the corresponding amino acids are shown in Table 1.

Second, we had to develop the general synthetic strategy. It was important to select an approach to generate an internucleotide link and to choose protective groups for internucleotide phosphates, terminal hydroxyls and 3′-phosphite. We decided to protect our target molecule **6** with the traditional protective groups for exocyclic aminogroups (benzoyl for adenine and cytosine and isobutyryl for guanine), for 5′-hydroxyl (dimethoxytrityl) and for phosphite residue (cyanoethyl).

**Table 1.** The structure of trinucleotide phosphoramidites synthesized. The corresponding amino acids are shown

| Codon | Amino acid | Codon | Amino acid |
|---|---|---|---|
| AAA | Lys | GAA | Glu |
| AAC | Asn | GAC | Asp |
| ACT | Thr | GCT | Ala |
| ATC | Ile | GGT | Gly |
| ATG | Met | GTT | Val |
| CAG | Gln | TAC | Tyr |
| CAT | His | TCT | Ser |
| CCG | Pro | TGC | Cys |
| CGT | Arg | TGG | Try |
| CTG | Leu | TTC | Phe |

The key substance in the synthesis of **6** is trinucleotide **5** with an unprotected 3′-hydroxyl function (Fig. 1). The main problem in the synthesis of compound **5** is to select a protective group for 3′-hydroxyl. This group has to be removed after synthesis of **5** before the synthesis of the phosphoramidite **6**. We were not able to use the acid-labile protection because of the presence of the acid-labile dimethoxytrityl residue. Furthermore, our internucleotide protective group and *N*-acyl strategy prevented the use of base labile protecting groups. The phenoxyacetyl (5) and tertbutyldimethylsilyl (14) residues were used for the 3′-protection in earlier studies, but the authors noted a high amount of side products during its cleavage.

We decided not to protect the 3′-hydroxyl at all. Such an approach was used earlier (15,16) during synthesis of oligonucleotides by the phosphotriester method. In this case we had to select the mildest way for internucleotide link formation to use the difference in the reactivity of primary and secondary hydroxyl functions. To us, the phosphotriester approach seemed the most appropriate. Besides, by using the phosphotriester method we could avoid additional oxidations that are necessary in phosphoramidite and *H*-phosphonate methods. We applied the traditional o-chlorophenyl residue for the protection of internucleotide phosphate and 1-mesitylenesulphonyl-3-nitro-1,2,4-triazole (MSNT) as the condensing reagent.

The synthetic route is outlined on Figure 1. The P-component **1** was allowed to react with the OH-component **2** in the presence of MSNT in pyridine. The 5′-*O*-dimethoxytrityl-dinucleotide **3a** was isolated by chromatography on silica-gel and treated with benzenesulphonic acid to remove the 5′-OH-protective group. The resulted 5′,3′-OH-dinucleotide was isolated by silica-gel chromatography in an average yield of 60–70%. The 5′-*O*-dimethoxytrityl-trinucleotide **5** was obtained by the same manner as dinucleotide **3a** in an 80–90% yield.

After the reaction of the P-component with a *N*-acyl nucleoside with free hydroxyls two dinucleotide isomers (**3a** and **3b**) and the trinucleotide **4** (a product of the reaction of **3** with excess of P-component) can be formed (Fig. 2). The trinucleotide **4** is much more lipophylic than dinucleotides **3** and can be easily separated by silica-gel chromatography. To separate dinucleotides **3** is more difficult because their chromatographic properties are rather

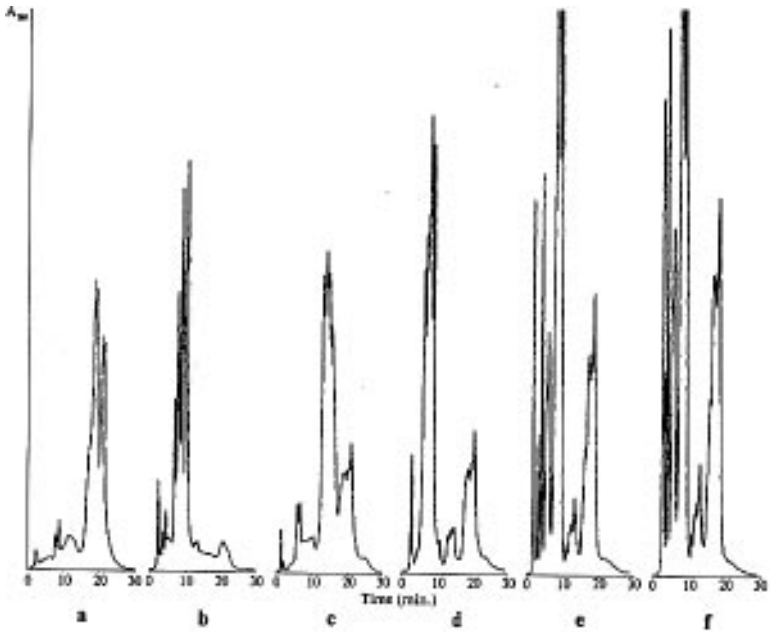


**Figure 3.** HPLC-chromatograms of test reactions for trinucleotide phosphoramidites **6**. (**a** and **c**) Chromatograms of trinucleotides AAC and GGG, (**b** and **d**) results of test reaction with these trinucleotides, (**e** and **f**) results of failed synthesis. See text for details.

similar. Moreover, both **3a** and **3b** form diastereomer pairs that may be separated during chromatography. This additionally hampers the assignment of peaks. Thus, the direct confirmation of the target product structure after each synthesis of dinucleotide and trinucleotide is unavoidable.

A confirmation of the di- and trinucleotides structure by spectrometric methods (NMR $^{31}$P, $^{1}$H and mass) is insufficient since neither NMR nor mass-spectra enable us to distinguish the 5′-3′- and the 3′-3′-isomers unambiguously. Therefore, we confirmed the structures of our di- and trinucleotides **3** and **5** by enzymatic digestion of fully deprotected derivatives. It is known that the snake venom phosphodiesterase (SPDE) splits the 5′-3′-phosphodiester link yielding 5′-mononucleotides and 5′-terminal nucleoside (if the oligonucleotide contains free 5′-hydroxyl). So, the HPLC analysis of the SPDE-hydrolysates provides complete information on the composition of **3a** and **5**. The SPDE-hydrolysates of **3b** and 3′-3′-isomers of **5** contain the starting material only, because the 3′-3′-isomers are stable to SPDE treatment.

We found that 5′-3′-isomers always move slightly slower on silica-gel than the corresponding 3′-3′-isomers and can be easily isolated as a pure mixture of diastereomers. Our optimized reaction conditions providing the maximal yield of the target products and the nucleotide composition of **3a** and **5** are listed in Tables 2 and 3.

The last step of synthesis of trinucleotide **6** was performed according to Nielsen *et al.* (7) and the target product was isolated by reversed-phase chromatography. Purity of the trinucleotides **6** was confirmed by reversed-phase HPLC as well. To interpret the chromatogram of **6** is difficult because each trinucleotide exists as a mixture of eight diastereomers. So we developed a simple test reaction that allows to distinguish peaks corresponding to 3′-phosphoroamidites. A solution of tetrazole in methanol was added to the solution of **6** in acetonitrile. The trinucleotides **6** were converted into corresponding *H*-phosphonates (shown by $^{31}$P-NMR spectra) under these conditions while all other possible impurities remained unchanged. Typical chromatograms are shown on Figure 3. The result of analysis of the AACphosphoroamidite is exemplified on Figure 3a and b. The figures show a complete reaction of this compound with tetrazole/methanol and no starting material (retention time 10–15 min) was left. The chromatograms of the GGG-phosphoramidite are presented on Figure 3c and d. This trinucleotide contained some impurities (retention time nearly 20 min) and was repurified. Finally, the results of a fully failed synthesis are shown on Figure 3e and f. No difference is observed between the starting material and the test reaction products. This preparation was discarded and resynthesized.

$^{1}$H-NMR spectra of all intermediates and target products were as expected. For short we do not present a full description of each spectrum. We describe instead the common features for each type of compounds—signals for protons of sugar skeleton, internucleotide protecting group, and so on, followed by the description of specific features of each compound—signals for protons of heterocyclic bases, exocyclic protective groups, etc. The overlapped signals (for instance, 2′-H of sugar moiety and CH of i-Bu-residue) were assigned by comparison of spectra of similar compounds (in this example—dimers AG and AA).

The compositions of **3a**, **5** and **6** were confirmed by mass-spectra. In all cases all expected signals were found. The strongest peak (m/e 302) corresponds to DTr-residue. Other main peaks (by decreasing of intensity): M+Na$^{+}$, m/e 240 (bzAde), 231 (i-buGua), 216 (bzCyt) and 126 (Thy, very weak). This picture is quite explainable in terms of stability of DTr-O and N-glycoside bonds. It is interesting that there is no difference between mass-spectra of **5** and **6** (of course, except of value M).

**Table 2.** Reaction conditions, isolated yields and nucleotide composition (content of 5′-terminal nucleoside is assigned to 1.0) of trinucleotides **3a**

| Dimer **3a** | OH-comp. (mmol) | P-comp. (mmol) | MSNT (mmol) | Yield (%) | Nucleotide composition pA | pC | pG | pT |
|---|---|---|---|---|---|---|---|---|
| AA | 15.7 | 18.0 | 27.0 | 70 | 1.1 | – | – | – |
| AC | 32.0 | 35.2 | 52.8 | 71 | – | 0.9 | – | – |
| AG | 7.0 | 7.8 | 11.6 | 66 | – | – | 1.1 | – |
| AT | 7.5 | 8.3 | 12.4 | 84 | – | – | – | 1.0 |
| CG | 7.5 | 8.3 | 12.4 | 68 | – | – | 1.1 | – |
| CT | 24.0 | 26.4 | 39.6 | 81 | – | – | – | 1.1 |
| GC | 7.5 | 8.3 | 12.4 | 59 | – | 0.9 | – | – |
| GG | 8.0 | 8.8 | 13.2 | 85 | – | – | 1.2 | – |
| GT | 18.0 | 19.8 | 29.7 | 65 | – | – | – | 1.1 |
| TC | 16.0 | 17.6 | 26.4 | 64 | – | 1.0 | – | – |
| TG | 20.0 | 22.0 | 33.0 | 61 | – | – | 1.1 | – |
| TT | 9.0 | 9.9 | 14.9 | 78 | – | – | – | 1.0 |

**Table 3.** Reaction conditions, isolated yields and nucleotide composition (content of 5′-terminal nucleoside is assigned to 1.0) of trinucleotides **5**

| Trimer **5** | OH-comp. (mmol) | P-comp. (mmol) | MSNT (mmol) | Yield (%) | Nucleotide composition pA | pC | pG | pT |
|---|---|---|---|---|---|---|---|---|
| AAA | 5.4 | 6.4 | 9.6 | 77 | 2.1 | – | – | – |
| AAC | 5.8 | 6.4 | 9.6 | 73 | 1.1 | 0.9 | – | – |
| ACT | 7.0 | 7.7 | 11.5 | 67 | – | 0.9 | – | 1.1 |
| ATC | 5.1 | 5.8 | 8.8 | 97 | – | 1.0 | – | 1.2 |
| ATG | 5.7 | 6.3 | 9.4 | 57 | – | – | 1.1 | 1.2 |
| CAG | 4.6 | 5.5 | 8.3 | 70 | 0.9 | – | 1.1 | – |
| CAT | 6.3 | 7.6 | 11.3 | 78 | 1.0 | – | – | 1.1 |
| CCG | 5.1 | 5.9 | 8.8 | 73 | – | 0.9 | 1.0 | – |
| CGT | 5.9 | 6.7 | 10.1 | 61 | – | – | 1.1 | 1.1 |
| CTG | 5.7 | 6.3 | 9.4 | 75 | – | – | 1.1 | 1.1 |
| GAA | 5.5 | 6.6 | 9.8 | 82 | 2.1 | – | – | – |
| GAC | 5.9 | 6.8 | 10.3 | 91 | 1.0 | 0.9 | – | – |
| GCT | 6.0 | 6.9 | 10.4 | 70 | – | 1.0 | – | 1.0 |
| GGT | 5.9 | 6.8 | 10.2 | 64 | – | – | 1.1 | 1.1 |
| GTT | 6.0 | 6.6 | 9.9 | 83 | – | – | – | 1.9 |
| TAC | 5.7 | 6.3 | 9.4 | 94 | 1.1 | 1.0 | – | – |
| TCT | 7.0 | 7.7 | 11.5 | 93 | – | 0.9 | – | 1.0 |
| TGC | 4.0 | 4.4 | 6.6 | 78 | – | 1.0 | 1.1 | – |
| TGG | 7.4 | 8.1 | 12.2 | 75 | – | – | 2.0 | – |
| TTC | 5.0 | 6.0 | 9.0 | 95 | – | 1.0 | | 1.1 |

Additionally all the phosphoramidites were tested by the $^{31}$P-NMR spectroscopy. The expected ratio between P(V) and P(III) signals was found in all cases. No signals in a range 5–6 p.p.m. (*H*-phosphonates) was observed. The component ratios, reaction yields and $^{31}$P-NMR data for **6** are listed in Table 4.

The trinucleotides **6** were used as synthons in oligonucleotide (library) synthesis. Every synthon was checked in a DNA-synthesis coupling to all four nucleobases in test sequences. The average coupling yield was calculated using data of quantitative gel analysis. Now every synthon got a reaction factor (RF, listed in Table 4) giving the most reactive one (trimer AAC) the factor 1. These factors compensate the differences in reactivity of trimers. For instance, the trimer AAC (RF = 1) is nearly twice more reactive than CAG (RF = 2.0). Therefore, the concentration of CAG in the mixture of trimers must be twice more than the concentration of AAC to obtain the equal incorporation of those trimers into the oligonucleotide. In that way we succeeded to create reaction mixtures of the trimers for incorporation of the amino acid codons into oligonucleotide in desired ratios. We cannot explain the abnormal high RF value for TGG-trimer. Apparently this preparation contains some impurities, which inhibit the reaction and cannot be detected by all analytical methods applied by us.

We used slightly modified standard cycles and procedures from the synthesizer producer—double or triple coupling and extended (2-fold) reaction times. Taking into account the reaction factors of trinucleotides, we are confident to receive oligonucleotide libraries with the even distribution of codons.

**Table 4.** Reaction conditions, isolated yields, reaction factors (RF) and $^{31}$P-NMR data of trinucleotides **6**

| Trimer **6** | Trimer **5** (mmol) | Tetrazole (mmol) | Reag. **7** (mmol) | Yield % | RF | $^{31}$P-NMR data δ P(V), p.p.m. | δ P(III), p.p.m. | P(V)/P(III) |
|---|---|---|---|---|---|---|---|---|
| AAA | 4.1 | 5.4 | 5.4 | 75 | 1.1 | –10.6 to –10.2 | 146.2–146.8 | 2.1:1 |
| AAC | 3.9 | 5.0 | 5.0 | 83 | 1.0 | –10.8 to –10.0 | 146.3–146.9 | 2.2:1 |
| ACT | 4.3 | 5.6 | 5.6 | 71 | 1.3 | –10.2 to –9.9 | 146.8–147.2 | 2.0:1 |
| ATC | 4.1 | 5.4 | 5.4 | 71 | 1.2 | –10.1 to –9.8 | 146.2–146.8 | 2.0:1 |
| ATG | 4.4 | 5.7 | 5.7 | 63 | 1.3 | –10.0 to –9.7 | 146.1–146.9 | 2.0:1 |
| CAG | 3.6 | 4.7 | 4.7 | 71 | 2.0 | –10.4 to –10.0 | 146.1–146.9 | 1.9:1 |
| CAT | 4.4 | 5.8 | 5.8 | 72 | 1.3 | –10.3 to –9.9 | 146.8–147.4 | 1.9:1 |
| CCG | 3.8 | 4.9 | 4.9 | 76 | 1.8 | –10.6 to –9.7 | 146.1–146.9 | 2.0:1 |
| CGT | 3.6 | 4.6 | 4.6 | 80 | 1.1 | –10.8 to –10.0 | 146.8–147.3 | 2.0:1 |
| CTG | 4.3 | 5.5 | 5.5 | 67 | 1.2 | –10.6 to –9.8 | 146.1–146.8 | 2.1:1 |
| GAA | 4.4 | 5.8 | 5.8 | 72 | 1.9 | –10.8 to –10.0 | 146.1–146.8 | 2.1:1 |
| GAC | 3.9 | 5.1 | 5.1 | 68 | 1.3 | –10.6 to –10.0 | 146.3–146.9 | 2.0:1 |
| GCT | 4.2 | 5.5 | 5.5 | 71 | 1.5 | –10.9 to –10.0 | 146.8–147.2 | 2.1:1 |
| GGT | 3.8 | 4.9 | 4.9 | 78 | 1.1 | –11.0 to –10.1 | 146.9–147.0 | 2.1:1 |
| GTT | 4.8 | 6.2 | 6.2 | 62 | 1.9 | –10.0 to –9.8 | 146.9–147.1 | 2.0:1 |
| TAC | 4.1 | 5.3 | 5.3 | 74 | 1.6 | –10.3 to –9.9 | 146.3–146.9 | 2.0:1 |
| TCT | 4.5 | 5.9 | 5.9 | 70 | 1.3 | –10.5 to –10.0 | 146.8–147.2 | 1.9:1 |
| TGC | 4.0 | 5.2 | 5.2 | 71 | 1.5 | –10.9 to –10.0 | 146.2–146.8 | 2.1:1 |
| TGG | 4.2 | 5.4 | 5.4 | 70 | 11.3 | –11.0 to –10.0 | 146.2–146.9 | 2.0:1 |
| TTC | 4.5 | 5.8 | 5.8 | 83 | 2.9 | –10.6 to –9.9 | 146.2–146.8 | 2.2:1 |

The following DNA-libraries have been synthesized:
(i) 5′-GAC-ACG-GCC-GTG-TAT-TAC-TGT-GTG-AGA-(Tri)$_8$-TGG-GGC-AAA-GGG-ACC-ACG-GTC-3′

Tri is an equivalent mixture of all 20 trimers based on the reaction factors. The yield of the first trimer coupling was 80% while subsequent trimers were coupled with 82–97% yield.
(ii) 5′-GCG-GTC-TAT-TTC-TGT-GCT-AGA-(Tnn)$_4$-GGA-TAC-TGG-GGC-CAA-GGG-ACT-3′

Tnn is an equivalent mixture of 19 trimers (trimer TGC was omitted). The reaction conditions were the same as mentioned above. In this case the coupling yields ranged from 71% (for the first trimer) up to 94–98% for trimers 2–4.

Obviously there is not a big difference in the coupling yields of trimers in comparison with three monomer couplings. Only the first coupling of a trimer mixture shows a significantly reduced coupling yield. At the moment we have no explanation for this effect but an additional solvent (dichloromethane) may be crucial. The trimer mixtures can be used in a normal synthesizer system without any problems. Cleavage of the protective groups does not appear to be a big problem when first performing it overnight at room temperature. If heating is performed too early, one can see broken DNA strands in a gel analysis.

We decided not to prove the quality of our libraries (number of different DNAs, amount of each DNA, etc.) by any chemical or physicochemical method. We have incorporated the libraries into biological (phage/bacterial) systems and now studies are in progress to test efficiency of these libraries for their use in applied molecular evolution.

## ACKNOWLEDGEMENTS

The authors would like to thank Drs T. Balashova, D. Dementieva and Yu Kozmin (Shemiakin & Ovchinnikov Institute of Bioorganic Chemistry, Moscow) for help with NMR and mass-spectra.

## REFERENCES

1 Kay,B.K.(1994) *Perspectives in Drug Discovery and Design*, **2**, 251–268.
2 Little,M., Breitling,F., Micheel,B. and Dubel,S. (1994) *Biotech. Adv.*, **12**, 539–555.
3 Pack,P. and Pluckthun,A. (1994) *Biochemistry*, **31**, 1579–1584.
4 Arkin,A.P. and Youvan,D.C. (1992) *BioTechnology*, **10**, 297–300.
5 Virnekas,B., Ge,L., Pluckthun,A., Schneider,K.V., Wellnhofer,G. and Moroney,S.E. (1994) *Nucleic Acids Res.*, **22**, 5600–5607.
6 Sproat,B.S. and Gait,M.J. (1984) In Gait,M.J. (ed.), *Oligonucleotide Synthesis—A Practical Approach*. IRL Press, Oxford, p. 98.
7 Nielsen,J. and Dahl,O. (1987) *Nucleic Acids Res.*, **15**, 3626.
8 Ikemura,T. (1981) *J. Mol. Biol.*, **146**, 1–21.
9 Smith,G.P. (1991) *Curr. Opin. Biotechnol.*, **2**, 668–673.
10 Griffiths,A.D., Williams,S.C., Hartley,O., Tomlinson,I.M., Waterhouse,P., Crosby,W.L., Kontermann,R.E., Jones,P.T., Low,N.M., Allison,T.J. *et al.* (1994) *EMBO J.*, **13**, 3245–3260.
11 Skerra,A. and Pluckthun,A. (1988) *Science*, **240**, 1038–1041.
12 Dubel,S., Breitling,F., Fuchs,P., Braunagel,M., Klewinghaus,I. and Little, M.(1993) *Gene*, **128**, 97–101.
13 Charbit,A., Molla,A., Saurin,W. and Hofnung,M. (1988) *Gene*, **70**, 181–189.
14 Lyttle,M.H., Napolitano,E.W., Calio,B.L. and Kauvar,L.M. (1995) *BioTechniques*, **19**, 274–281.
15 Catlin,J.C. and Cramer,F. (1973) *J. Org. Chem.*, **38**, 245–248.
16 Agarwal,K. and Riftina,F. (1978) *Nucleic Acids Res.*, **5**, 2809–2823.