Efficient and flexible access to fully protected trinucleotides suitable for DNA synthesis by automated phosphoramidite chemistry

Andrea Zehl,[a] Antje Starke,[a] Dieter Cech,[a] Thomas Hartsch,[b] Rainer Merkl[b] and Hans-Joachim Fritz*[b]†

[a] *Institut für Chemie, Humboldt-Universität zu Berlin, Hessische Strasse 1-2, D-10099 Berlin, Germany*
[b] *Institut für Molekulare Genetik, Georg-August-Universität Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany*

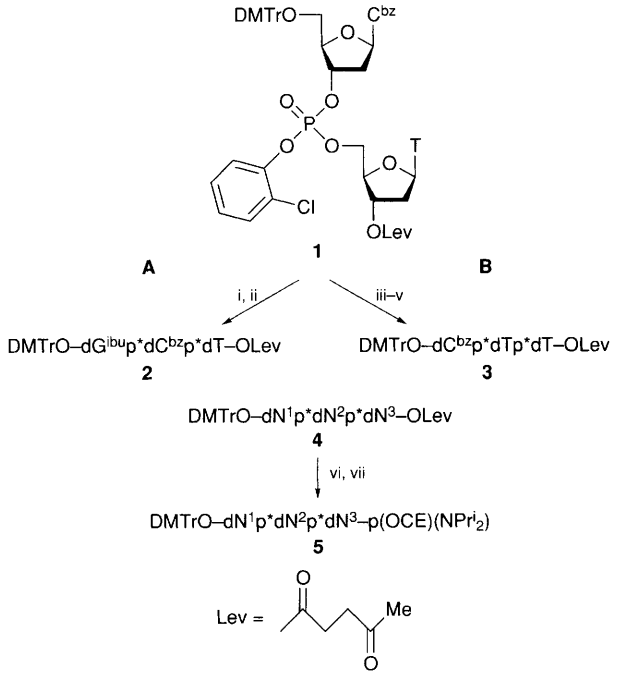
An efficient synthetic approach to protected trinucleotide phosphoramidites suitable for codon-by-codon synthesis of structural gene combinatorial libraries is described; the procedure rests on the use of a pair of orthogonal protecting groups for the two oligomer termini and offers flexibility in the choice of chain elongation direction going from dimer to trimer.

Protected trinucleotides suitable for incorporation into synthetic DNA are necessary reagents in the controlled, codon-by-codon construction of combinatorial libraries of structural genes. All the methods of preparing these building blocks that have been published to date,[1–3] however, suffer from one or more limitations. We describe a method that gives easy and flexible access to trinucleotide blocks capable of undergoing coupling reactions by the phosphoramidite method. In particular, our method is characterized by use of a pair of orthogonal protecting groups for the 5′- and the 3′-terminus and, as a consequence, flexibility with respect to direction of chain elongation and optimal use of synthetic intermediates. The approach is outlined in Scheme 1. For convenient handling of soluble intermediates on a large scale, phosphotriester chemistry with the *ortho*-chlorophenyl (*o*-ClPh) group as the phosphate protecting group was applied to all internucleotidic links to be formed in solution, (*i.e.* the internal phosphates of a trinucleotide diphosphate intermediate),[4–6] whereas a standard phosphoramidite function was introduced to the 3′-terminus of each building block for assembly of oligonucleotides on solid support.[7]

With the commonly-used dimethoxytrityl (DMTr) group, used for 5′-OH protection, kept constant, optimal choice of a 3′-OH protecting group becomes the key feature of the approach. Virnekäs *et al.*[1] used the phenoxyacetyl group for 3′-*O*-protection and methanolic ammonia for removal, conditions known to give rise to breakage of internucleotidic triester linkages[8] irrespective of their P^{III} or P^{V} nature. The *tert*-butyldimethylsilyl (TBDMS) group[9] can easily be removed by fluoride ion; this reaction, however, is accompanied by cleavage of phosphotriester bonds. Acidic removal of TBDMS, as practised by Lyttle *et al.*,[2] also has severe drawbacks, such as concomitant loss of the 5′-DMTr group. Finally, Ono *et al.*[3] have described the synthesis of 3′-unprotected trinucleoside diphosphates by reaction of 3′,5′-unprotected nucleosides with 5′-dimethoxytritylated dinucleoside diphosphates. The convenience of relinquishing hydroxy group protection altogether has its price; inevitably there has to be some undesired 3′-*O*-phosphorylation of the nucleoside.

We found the levulinyl (Lev) group to meet all requirements for 3′-*O*-protection. It is removed by hydrazine, in pyridine–acetic acid; a treatment that not only leaves the 5′-DMTr group intact but also poses no risk to internucleotidic linkages or glycosidic bonds.[10,11] Furthermore, the orthogonality of DMTr and Lev allows fully protected dinucleoside monophosphates **1** (Scheme 1) to be elongated to trinucleotides in either the 3′ to 5′ (route **A**) or the 5′ to 3′ direction (route **B**). Scheme 1 outlines the procedure for the synthesis of two trinucleotide blocks (fully protected GCT and CTT) using the common intermediate **1**.

Following the procedures described in detail below, we have synthesized the trinucleotides summarized in Fig. 1.



Scheme 1 *Reagents and conditions*: i, Cl_3CCO_2H (0.39 M) in $CHCl_3$, 0 °C, 10 min; ii, 5′-DMTr-dG^{ibu}-3′-OP(O)$_2$OC$_6$H$_4$Cl (1.2 equiv., Et_3NH salt), *N*-methylimidazole (9.5 equiv.) 2,4,6-triisopropylbenzene sulfonyl chloride (TIPBSCl) (2.97 equiv.), room temp., 15 min, pyridine; iii, hydrazine hydrate (0.5 M; 10 ml) in pyridine–acetic acid (3 : 2, v/v), room temp., 5 min; iv, triazole (6 mmol), 2-chlorophenyldichlorophosphate (2.25 mmol), NEt_3 (5.25 mmol), THF (15 ml), then add **1** (1 mmol), room temp., 1 h, then triethylammonium hydrogen carbonate (1 M; 15 ml; pH 8.0); v, 3′-OLevT (0.83 equiv.), *N*-methylimidazole (7.92 equiv.), TIPBSCl (2.47 equiv.), room temp., 15 min, pyridine; vi, as for iii, with N = A^{bz} (*N*^{6}-benzoyladenosine residue), C^{bz} (*N*^{4}-benzoylcytidine residue), G^{ibu} (*N*^{2}-isobutyrylguanosine residue) or T (thymidine residue); vii, MeCN–CH_2Cl_2 (1 : 1, v/v), 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (2 equiv.), diisopropylethylamine (5 equiv.) room temp., 2 h

5′TCXYZT3′

XYZ	XYZ	XYZ
TTT	TTC	AAC
CTT	ATC	CAC
GTT	ATG	GAC
GGG	CTG	TAC
TGG	GGT	TCT
GAA	CGT	GCT

Fig. 1 Automated synthesis of model hexanucleotides. Trinucleotide blocks (XYZ) were introduced *via* the sample injector module; monomers through the standard phosphoramidite ports of the Pharmacia Gene Assembler. 'Pharmacia Primer Support T' cartridges were used as solid support.

Fully protected dinucleoside monophosphates of type **1** were prepared using standard methods of phosphotriester oligonucleotide synthesis.[12] Addition of another nucleotide residue to either the 5′- or the 3′-terminus of **1** was accomplished *via* routes **A** or **B**, respectively.[12] Following route **A**, the DMTr group was removed by acid [step (i)] and the 5′-*O*-deprotected dinucleoside monophosphate was condensed with a 5′-*O*-dimethoxytritylated nucleoside-3′-phosphate carrying the *o*-ClPh phosphate protecting group [step(ii)] to yield fully protected trinucleoside diphosphates of type **2**. Alternatively, following route **B**, the Lev group was removed from the 3′-*O*-terminus of the type **1** compound [step (iii)], followed by phosphorylation [step (iv)] and condensation to a 3′-*O*-Lev nucleoside carrying a free 5′-hydroxy function [step (v)]. This gives fully protected trinucleoside diphosphates **3** which have the same general structure as those of type **2**. With both routes, trinucleotide yields are in the range of 70–88% relative to the coupling component used in sub-equivalent amount.

The option of elongating a given fully protected dinucleoside monophosphate in either direction adds flexibility to the approach in that the number of dimers necessary to synthesize a certain set of trinucleotides can be reduced and/or notoriously inefficient G to G couplings can be avoided in the more expensive step going from dimer to trimer (except, of course, in the case of GGG). The fully protected trinucleoside diphosphates **4** were converted to the corresponding 3′-phosphoramidites **5** by removing the Lev group[12] [step (vi)], followed by phosphitylation according to standard procedures [step (vii)]. Type **5** compounds are the target coupling blocks; their yields were in the range of 85–95%, relative to starting compounds **4**.

Next, the trinucleotide blocks **5** were tested for their ability to be incorporated into oligonucleotide chains by phosphoramidite chemistry. Eventually, defined mixtures of trinucleotides are to be used in a single synthesis step in order to prepare combinatorial libraries. Such a scheme is not supported by conventional DNA synthesizers that have only a few reagent ports for phosphoramidites. Since, however, it seemed highly desirable to use the option of machine-aided synthesis, we first constructed a suitable DNA synthesizer by combining a Pharmacia GeneAssembler with a computer-controlled chromatographic sample injector (Pharmacia model 2157 autosampler) and making adjustments to the control software (details to be described elsewhere). In this configuration, chain elongation components needed in individual steps of a synthesis can be loaded onto the sample injector in a row of vials, from which they are recruited for synthesis by activating the injector at the appropriate time.

The hexanucleotides illustrated in Fig. 1 were synthesized. After a first survey of reaction conditions, the coupling reaction time for trinucleotide blocks was set to 15 min (as compared to the routinely used 2 min); all other synthesis parameters were kept constant. Under these conditions, efficiencies of triblock coupling reached 99%, with considerable variation depending on the particular batch. Not surprisingly, however, type **5** trinucleotide block GGG had a coupling yield of only 69%. In the next set of experiments, mixtures of trinucleotide coupling blocks were used, total trinucleotide amount added always being 50 μl of a 0.1 M solution which corresponds to a 25-fold excess over the 0.2 μmol of chain termini bound to the solid support. The resulting hexamer mixtures were analysed by reversed phase HPLC after partial deprotection (5′-DMTr group on). Three triblocks that individually showed similar coupling yields (TTC, AAC and CAC) were represented in the hexamer mixture in comparable quantities (39, 29 and 31%, respectively) when present in the coupling reaction in equimolar mixture. A likewise equimolar mixture of TTT and GGG blocks resulted in 66% total coupling yield and a representation ratio in the hexamer mixture of 61 : 39. Significantly, this ratio could be brought to roughly 1 : 1 when the amount of GGG was doubled relative to TTT and could be reversed by a further doubling of the amount of GGG.

In order to check triblock coupling yields in the synthesis of longer oligonucleotides and with a variety of nucleotide residues as the 5′-OH component of the reaction, the following compounds were synthesized (yield of triblock couplings given in parentheses; residues not in bold were added as monomers): 5′-TCCGA**CTT**CTATGAC-3′ (93%); 5′-TCCGC**TTT**GTATGAC-3′ (99%), 5′-TCCGT**GTT**ATATGAC-3′ (88%), 5′-TCCGT**GGG**CTATGAA-3′ (66%). Final products were compared by gel electrophoresis and reversed phase HPLC to the respective oligonucleotides identical in sequence but synthesized separately from monomers only; no differences in product identity nor homogeneity were detected.

Combined use of monomeric and trimeric coupling units as described here is clearly sufficient for the combinatorial synthesis of gene libraries with entirely controlled degeneracy at a number of selected, individual codon positions. While this covers most applications in evolutionary protein engineering, the eventual synthesis of long polynucleotides exclusively from trimers will necessitate further improvement of coupling yields. Work along these lines is in progress.

This work was supported by the German Federal Minister of Education and Research (BMBF) and by Fonds der Chemischen Industrie.

Footnote

† E-mail: hfritz@Uni-Molgen.gwdg.de

References

1 B. Virnekäs, L. Ge, A. Plückthun, K. C. Schneider, G. Wellnhofer and S. E. Moroney, *Nucleic Acids Res.*, 1994, **22**, 5600.
2 M. H. Lyttle, E. W. Napolitano, B. L. Calio and L. M. Kauvar, *BioTechniques*, 1995, **19**, 274.
3 A. Ono, A. Matsuda, J. Zhao and D. V. Santi, *Nucleic Acids Res.*, 1995, **23**, 4677.
4 C. B. Reese, *Colloq. Int. CNRS*, 1970, **182**, 319.
5 K. Itakura, C. P. Bahl, N. Katagiri, J. J. Michniewicz, R. H. Wightman and S. A. Narang, *Can. J. Chem.*, 1973, **51**, 3649.
6 V. A. Efimov, S. V. Reverdatto and O. G. Chakhmakhcheva, *Nucleic Acids Res.*, 1982, **10**, 6675.
7 J. Nielsen, M. Taagaard, J. E. Marugg, J. H. van Boom and O. Dahl, *Nucleic Acids Res.*, 1986, **14**, 7391.
8 M. J. Gait, in *Oligonucleotide Synthesis: a Practical Approach*, ed. M. J. Gait, IRL Press, Oxford, 1984, p. 13.
9 K. K. Ogilvie, *Can. J. Chem.*, 1973, **51**, 3799.
10 J. H. van Boom and P. M. J. Burgers, *Recl. Trav. Chim. Pays-Bas.*, 1978, **97**, 73.
11 K. K. Ogilvie and M. J. Nemer, *Can. J. Chem.*, 1980, **58**, 1389.
12 M. J. Gait, H. W. D. Matthes, M. Singh, B. S. Sproat and R. C. Titmas, in *Chemical and Enzymatic Synthesis of Gene Fragments*, ed., H. G. Gassen and A. Lang, Verlag Chemie, Weinheim, 1982, pp. 31–37.

Received, 19th August 1996; Com. 6/05786A