Conversion of Single-Stranded Oligonucleotides into Cloned Duplexes and its Consecutive Application to Short Artificial Genes

Imre Cserpan,[a] Miklos Kalman,[a] Marie-Louise Tjörnhammar[b] and Andras Simoncsits[a,b]

[a] Institute of Genetics, Biological Research Center, Hungarian Academy of Sciences, H-6701 Szeged, Hungary and [b] Department of Biochemistry, Arrhenius Laboratories for Natural Sciences, Stockholm University, S-106 91 Stockholm, Sweden

Cserpan, I., Kalman, M., Tjörnhammar, M.-L., Simoncsits, A., 1991. Conversion of Single-Stranded Oligonucleotides into Cloned Duplexes and its Consecutive Application to Short Artificial Genes. – Acta Chem. Scand. 45: 265–272.

A general method to convert single-stranded, chemically synthesized oligonucleotides into cloned duplexes is described. Oligonucleotides supplied with 3′-terminal extensions that are complementary to 3′-protruding ends obtained by certain restriction enzymes can be cloned either directly or with the help of an adapter molecule into double-stranded vectors. Two methods have also been developed for consecutive cloning applications. According to these methods, the synthetic oligonucleotides (and their enzymatically prepared complementary strands) are joined, one after the other, inside a cloning vector, each joining requiring one cloning step. Synthetic genes are thus built up from oligonucleotides corresponding to only one strand of the DNA. The sequential assembly of the cloned duplex takes place in the 5′ to 3′ direction. Each oligonucleotide is supplied with a four-nucleotide-long 3′-terminal extension, but this sequence is eliminated when the joining takes place, leaving no limiting sequence between the oligonucleotides. The two consecutive cloning methods, the adapter and the polycloning site methods, are illustrated by the assembly of short artificial genes.

Total synthesis of genes has now become a widespread technique.[1] The classical gene assembly strategy[2] which has, with modifications, been successfully used for longer and longer genes[3,4] requires the chemical synthesis of both strands of the DNA. To reduce the amount of the chemical synthesis, a method has been developed in which long, synthetic oligonucleotides annealed at their 3′-terminal complementary sequences form a partial duplex which is then converted into a full duplex by DNA polymerase.[5,6] This duplex could then be cloned in an oriented way by utilizing the preformed restriction recognition sequences at the 5′-terminal regions of the starting oligonucleotides.[6,7] The reduction in the amount of the chemical work depends strongly on the lengths of the starting oligonucleotides, considering that the overlapping region (8–10 nucleotides) is synthesized as a duplex while other regions do not appear in the cloned sequence (the upstream sequences of the preformed restriction sites which are included only to facilitate the restriction cleavage).

To achieve further reductions of the chemical work involved in total gene synthesis, we have been developing single-stranded cloning methods which are different from the existing ones. Previous methods have utilized both the 5′- and the 3′-extensions of the synthetic oligonucleotide, which are complementary to the cohesive vector ends generated by a pair of restriction enzymes.[8,9] Ligation of both ends of the oligonucleotide to the linear vector resulted in a gapped duplex which was directly transformed into *Escherichia coli* for the gap repair to take place *in vivo*. The gap repair method has further been developed by utilizing bridging oligonucleotides[10,11] or by incorporating long arms to the ends of the oligonucleotide to be cloned.[12]

Our major aim has been to develop methods which are suitable not only for the cloning of the single-stranded oligonucleotides, but also for the cyclic application of this process to result in cloned genes. These methods include consecutive cloning and joining of single-stranded oligonucleotides and have been applied for both short[13,14] and long (over 1700 base pairs) synthetic genes.[15] In this work we describe the single-stranded cloning method we use, as well as its extended versions (the adapter and the polycloning site methods) developed for consecutive applications.

Experimental

General. Restriction endonucleases and T4 DNA ligase were from New England Biolabs, T4 polynucleotide kinase and *E. coli* DNA polymerase I (Klenow fragment) were from Boehringer. Radiochemicals were purchased from Amersham. Chemicals for the phosphate triester method were obtained from Chruachem, Scotland, except for the dimer building blocks which were synthesized.[16] Phosphoramidite reagents were from Pharmacia. Vectors M13mp10,[17] M13mp18[18] and pUC8[19] were from Pharmacia and were propagated in JM101 *E. coli*.[20] Transformations were carried out according to Hanahan.[21] Standard enzymatic reactions and cloning techniques were performed as described.[22]

Synthetic oligonucleotides. Sequences are written in the 5′ to 3′ direction. The 45-mer CATGGGCATCGTTGAACAGTGTTGTACTTCTATCTGCTCTCTGCA (human insulin A chain first oligonucleotide, IA1), the 39-mer CATGTTTGTTAACCAGCACCTGTGCGGCTCTCACCTGCA (human insulin B chain first oligonucleotide, IB1), the 35-mer TTTACCAGCTTGAGAACTACTGTAACTAGCCTGCA (human insulin A chain second oligonucleotide, IA2), the 67-mer TGGTTGAAGCTCTGTACCTGGTTTGCGGCGAACCTGGCTTCTTCTACACTCCGAAAACTTAGCTGCA (human insulin B chain second oligonucleotide, IB2) and the PstI-HindIII adapter

5′ GCGGGGGGCA 3′
3′ GACGTCGCCCCCCGTTCGA 5′

were prepared by the phosphate triester methodology on a manual DNA bench synthesizer (Omnifit) using monomer or/and dimer building blocks as described.[23] The 55-mer ATGCACTCTGACGCTGTTTTCACTGACAACTACACTCGTCTGCGTAAACAGGTAC (vasoactive intestinal polypeptide first oligonucleotide, VIP1), the 45-mer CTGGCTGTTAAAAAGTACCTGAACTCTATCCTGAACGGCTGAGCT (VIP second oligonucleotide, VIP2) and the CAGTTTCTTTATGGCAGGGT sequencing primer located upstream of the lacZ ClaI site were prepared by the phosphoramidite method on an automated synthesizer (Pharmacia Gene Assembler). Purifications were performed on preparative polyacrylamide gels containing 8 M urea.

Ligation of the oligonucleotides to the PstI-HindIII adapter. The 5′-[^{32}P]phosphorylated adapter strands (20 pmol each) were mixed in 10 µl of water and left at 60 °C for 30 min, then 5′-[^{32}P]phosphorylated IA2 or IB2 oligonucleotide (33 pmol in 5 µl of water) was added. The mixture was kept at 37 °C for 10 min, cooled to 16 °C and was made up to 20 µl volume containing 50 mM Tris/HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol and 1 mM ATP (ligase buffer). One hundred units of T4 DNA ligase (0.25 µl) were then added and the mixture was kept at 16 °C for 16 h, then treated at 60 °C for 10 min. After ethanol precipitation and drying, the pellet was dissolved in 20 µl of buffer (50 mM NaCl, 10 mM Tris/HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM dithiothreitol) and 20 units of HindIII were added at 37 °C over 2 h. The oligonucleotide–adapter complex was separated from the other oligonucleotide components by non-denaturing gel electrophoresis on a 10 % polyacrylamide gel and isolated after autoradiography, gel elution and ethanol precipitation.

Cloning of single-stranded oligonucleotides. The 5′-phosphorylated oligonucleotide (0.02–5 pmol) was reacted with double cleaved, gel-purified vector (0.02 pmol) in 10 µl reaction mixture containing 50 mM Tris/HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP (ligase buffer) and 80 units of T4 DNA ligase (0.2 µl) at 16 °C for 4–16 h. The reaction mixture was heated to 60 °C for 5 min, cooled to room temperature and 1 µl of 1 mM dNTP (an equimolar mixture of all four deoxynucleoside 5′-triphosphates) and 1 µl of 0.5 units µl^{-1} Klenow polymerase were added. After being allowed to stand at room temperature for 30 min, the mixture was heat treated (60 °C, 15 min) again, cooled to 16 °C and 8 µl of ligase buffer and 200 units of T4 DNA ligase (0.5 µl) were added. The second ligation was performed for 6–20 h at 16 °C. Aliquots of the reaction mixture were transformed into JM101 cells and plated onto TY-plates in soft agar containing isopropyl β-D-thiogalactopyranoside and 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (M13-derivatives) or onto LB-plates containing 100 µg ml^{-1} ampicillin (plasmid vectors). The oligonucleotide–adapter complexes were cloned essentially as described above.

Nucleotide sequencing. Dideoxynucleotide sequencing was performed on either a single-stranded[24,25] or denatured plasmid[26,27] template.

pPEX E. coli expression vector. This was constructed from a *rac* fusion promoter vector[28] containing the *N*-terminal 280 amino acids coding region of the lacZ gene up to the ClaI site[29] followed by a synthetic polycloning region containing unique restriction sites for BamHI, KpnI, SstI, ApaI and EcoRI enzymes.[14]

Results

Principle of the single-stranded cloning method. The cloning vector is cleaved with two different restriction enzymes to give it the required 3′-protruding cohesive end and usually a 5′-protruding cohesive end. The oligonucleotide to be cloned contains a four-nucelotide-long 3′-terminal sequence which is complementary to the 3′-protruding end of the cleaved vector. The 5′-phosphorylated oligonucleotide is ligated, through its 3′-terminal sequence, to the vector, so that the ligation takes place only at the 3′-terminal complementary regions but not at the 5′-end of the oligonucleotide. The 5′-terminal single-stranded extensions are then converted into duplexes with Klenow polymerase in the presence of all four deoxynucleoside 5′-triphosphates (dNTP mixtures). A linear vector possessing blunt ends is obtained which is then recircularized by T4 DNA ligase, followed by transformation of the reaction mixture into *E. coli* (Fig. 1).

Selection of the recombinants can utilize phenotypic changes caused by the insertion of the oligonucleotide into the vector. The M13mp phage and the pUC plasmid vectors are well suited for such selection as they contain the required unique restriction sites inside a region which codes for the *N*-terminal part of the β-galactosidase.[17–20] Very simple selection (blue to white colour change of the recombinant clones) is possible when the oligonucleotide insertion disrupts the coding information in the β-galactosidase

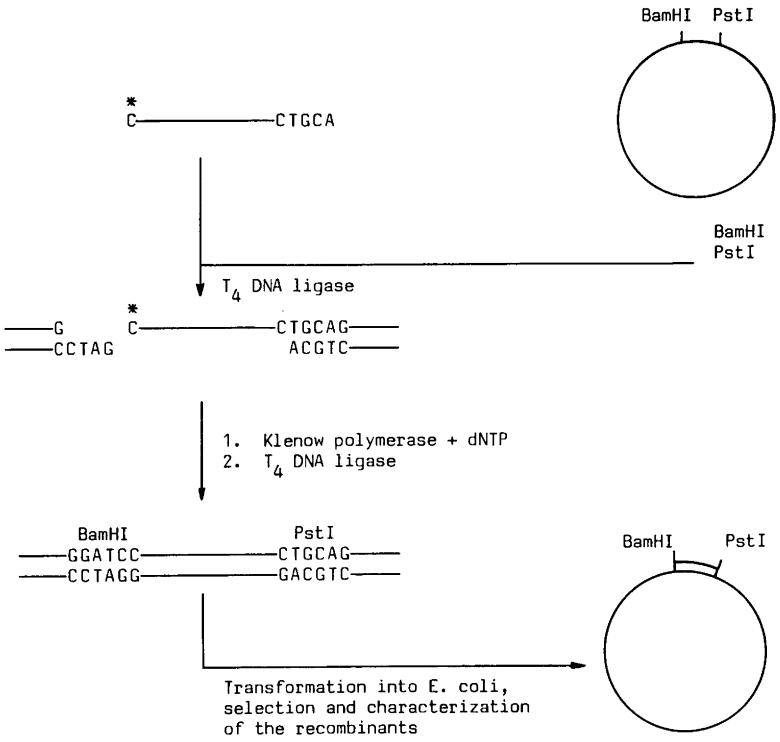

Fig. 1. General strategy for cloning single-stranded oligonucleotides. The oligonucleotide to be cloned is 5′-phosphorylated (asterisk over the 5′-terminal base). The open box in the recombinant represents the cloned oligonucleotide.

portion.[20] A more general selection possibility is the hybridization of the recombinant clones with the corresponding labelled oligonucleotide probe. The selected clones are further analysed by nucleotide sequencing. Preliminary screening of large number of clones with T-track analysis[24] may reduce the number of clones to be sequenced.

Applications of the single-stranded cloning method. The above principle is exemplified by the successful cloning of three oligonucleotides into three different cloning vectors, using different selection possibilities. These oligonucleotides represent the 5′-terminal part of short artificial genes.

As the first example, the 55-mer VIP1 oligonucleotide possessing a 3′-terminal extension characteristic of the KpnI site was chosen to optimize the ligation and cloning conditions and to analyse the recombinants. In combination with BamHI–KpnI cleaved M13mp18 an easy, blue-to-white selection is possible because the insertion of this oligonucleotide causes a frameshift in the alpha-peptide coding region. We wanted to verify the reliability of this selection, therefore reaction mixtures with different oligomer:vector ratios were made up and subjected to the steps of the cloning procedure. The results of seven different experiments are summarized in Table 1. The frequency of the white plaques, expressed as a percentage of the total plaques obtained, was increased by increasing the oligonucleotide : vector molar ratio from 1 to 50, but over 50 no further increase was obtained. This is in accord with the results of a parallel experiment in which a 5′-[^{32}P]phosphorylated oligonucleotide was used and its ligation to the vector was analysed by means of agarose gel electrophoresis followed by autoradiography (not shown). A sharp increase in the yield of the oligonucleotide-ligated vector was observed up to a 50 molar ratio but not over this range.

White recombinants obtained in each of the above experiments were subjected to T-track analysis. Clones showing the expected T-track were found infrequently in experiments A and B where the oligonucleotide was used in a low molar excess (one- and five-fold, respectively). Using a

Table 1. Cloning of the 55-mer VIP1 oligonucleotide into M13mp18 vector. The vector was cleaved with BamHI and KpnI, except for experiment G, when XbaI–KpnI cleaved M13mp18 was used. Experiments were performed as described in the Experimental except for the following: experiment F, short (1 h) first ligation followed by addition of the components of the second and third step at 16 °C then incubation at 16 °C for 16 h.

Experiment	Oligomer: vector ratio	White plaque (%)	Correct T-track	Correct sequence
A	1:1	3	0/6	–
B	5:1	6	2/6	2/2
C	15:1	10	3/4	2/3
D	50:1	16	23/30	10/11
E	250:1	15	5/6	2/2
F	50:1	15	5/6	1/1
G	50:1	12	9/12	4/4

higher molar excess, the yield of the recombinants with the correct T-track was increased to 75–83 %. Recombinants selected by the colour selection followed by T-track analysis were finally characterized by sequencing. Of the 23 clones sequenced, 21 contained the correct oligonucleotide sequence while two were found to contain one or two base substitutions. According to these results the probability of finding the correct, cloned sequence among the doubly selected recombinants was at least 90 %.

To understand the nature of the side reactions leading to aberrant clones and to increase the yield of the proper recombinants, blue plaques as well as white plaques showing incorrect T-tracks were taken for further characterization. The blue clones tested proved to be either the starting M13mp18 or its nine-base-pair deletion derivative. The high proportion of M13mp18 could be explained by incomplete double-cleavage followed by religation, while the deletion derivative could be obtained from the non-ligated, BamHI–KpnI cleaved vector when both the fill-in repair (BamHI end) and the 3′–5′ exonuclease (KpnI end) activities of the Klenow polymerase were exerted forming blunt ends, which were religated. Since this religation resulted in a BamHI site, it was realized that the blue background could be eliminated or decreased by BamHI cleavage performed before transformation. Indeed, the yield of the white plaques in experiments C–F (Table 1) could be increased from 10–15 % to 80–90 % (data not shown). Although extensive analysis of the white recombinants showing unexpected T-tracks was not performed, some classes of them could be identified. Compared with M13mp18 vector, four-base-pair deletions or insertions were frequently found. These could originate from a portion of the linear vector cleaved by only one of the enzymes, by blunt end formation followed by religation. Short, unspecified vector deletions without or with oligonucleotide insertion were also found. In some cases the cloned oligonucleotide region itself contained randomly located, one-base-pair deletions. Insertion of the oligonucleotide in the opposite orientation was also observed, which could be obtained from a vector cleaved by KpnI only. These results suggest that the efficiency of this cloning procedure depends strongly on the success of the double cleavage. More distantly spaced restriction sites in the vector would ensure more efficient double cleavage, therefore both the blue vector background and a part of the unwanted white recombinants could be reduced. In spite of the large variety of the aberrant clones selected by the colour change, their occurrence was not frequent and the selection was highly reliable when the oligonucleotide was used in a large molar excess. The expected VIP1 oligonucleotide sequence located between the utilized restriction sites was easily found among the recombinants. This clone (M13mp18VIP1) contains the 55-mer model oligonucleotide which is the 5′-terminal part of an artificial gene coding for a precursor of the human vasoactive intestinal polypeptide (VIP).

The second example is of cloning when colour selection for the recombinants was not possible. Insertion of the 45-mer oligonucleotide IA1 into BamHI and PstI-cleaved M13mp10 vector is expected to maintain the lac^+ phenotype. When the cloning experiment was performed, both blue and white plaques were obtained. The white plaques were probably obtained from the vector portion which had been cleaved only once and converted into a blunt end as discussed above. To find the desired recombinant, 50 blue plaques were taken for C-track analysis. Of these, 24 (48 %) were found to contain a 45-mer oligonucleotide insert, while the rest were shown to be either M13mp10 or its 18-base-pair deletion derivative. 14 clones selected by C-track analysis were sequenced and in eight cases (57 %) the correct oligonucleotide was found, while the rest contained point mutation(s). The correct clone (M13mp10IA1) contains the 5′-terminal part of a gene coding for the human insulin A chain.

In the third example the 39-mer oligonucleotide IB1 was cloned into BamHI and PstI-cleaved pUC8 vector, and the recombinants were selected by colony hybridization. Experiments performed with different oligomer : vector ratios followed by selection showed results similar to those described for the VIP1 oligonucleotide: the proportion of the positive clones sharply increased with increasing oligonucleotide : vector ratios and it reached an optimum at ratios in the range 50–200 (data not shown). Sequencing showed that seven of 12 selected clones (58 %) contained the correct IB1 sequence. The pUC8IB1 clone obtained contains the 5′-terminal part of an artificial gene coding for the human insulin B chain.

Principle of the consecutive cloning methods. The consecutive application of the single-stranded cloning method is shown schematically in Fig. 2. The first oligonucleotide is cloned, as described above, between sites A and B, and the recombinant obtained is used as a cloning vector in the next step. The recombinant vector is cleaved at the 3′-terminus of the cloned oligonucleotide (site B), exposing the 3′-terminal extension of the cloned oligonucleotide as a 3′-protruding end, and at a downstream position with a different enzyme resulting in a 5′- or 3′-protruding end [site C in Fig. 2(a) and site D in Fig. 2(b), respectively]. The downstream cohesive end (C or D) serves as the ligation site of the next oligonucleotide.

In the *adapter approach* [Fig. 2(a)], the downstream cohesive end (C) is usually a 5′-protruding one and the oligonucleotide can be ligated to the B–C cleaved vector through a universal adapter molecule having the corresponding cohesive ends for both the second oligonucleotide 3′-terminal extension (site B) and the downstream C site (B–C adapter). The oligonucleotide is ligated to the B–C adapter at site B, then the partial duplex, containing a 5′-terminal single-stranded region, a regenerated B restriction site and a 3′-terminal cohesive end (C), is ligated to the B–C cleaved vector so that the further ligation takes place only at the C cohesive ends. The linear vector obtained is treated with Klenow polymerase in the presence of dNTP mixture to fill in the 5′-extension representing the second

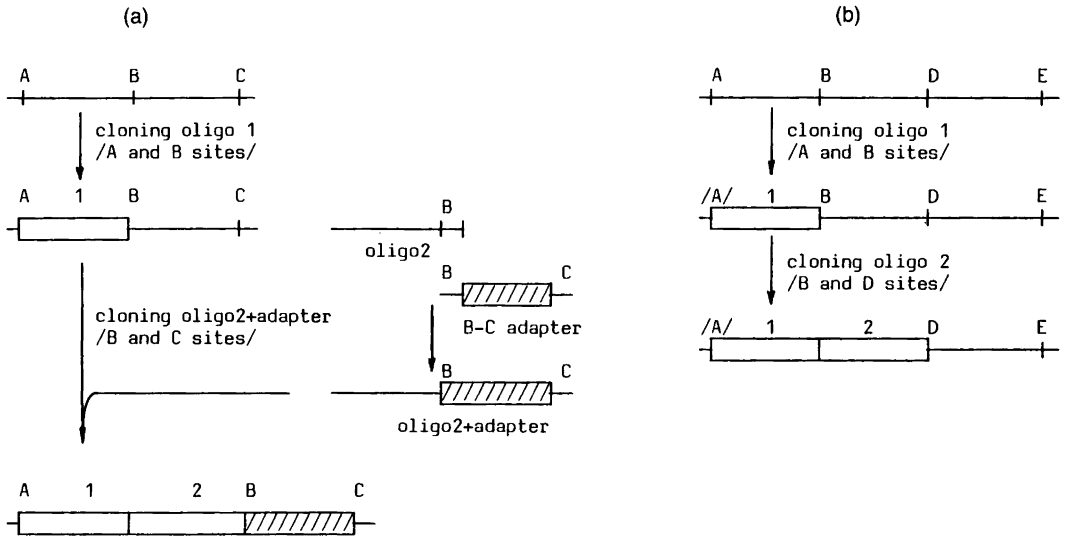

Fig. 2. Scheme of the consecutive cloning methods: (a) adapter method; (b) polycloning site method. Open boxes, the cloned oligonucleotides; shaded box, adapter sequences.

oligonucleotide and to remove the 3′-terminal single-stranded extension (3′-protruding B end). As a result of the conversions, the vector contains blunt ends corresponding to the 3′-terminus of the first oligonucleotide (without its original 3′-terminal extension) and to the 5′-terminus of the second. Recircularization of the vector by blunt-end ligation results in the joining of the two oligonucleotides. The reaction mixture is transformed into *E. coli* and the recombinants can be screened either by utilizing a phenotypic change caused by the insertion of the second oligonucleotide + adapter or by hybridization with the second, labelled oligonucleotide.

In the *polycloning site approach* [Fig. 2(b)], the cloning vector contains an array of restriction sites (B, D and E) resulting in different 3′-protruding ends. The recombinant vector containing the first-cloned oligonucleotide between sites A and B is cleaved with enzymes B and D and the second oligonucleotide, supplied with a 3′-terminal extension characteristic for site D, is ligated to this latter site of the linear vector. Further cloning steps are performed as described above.

The examples described in this work are short synthetic genes obtained by the joining of only two oligonucleotides, but the cyclic application should also permit the assembly of longer synthetic genes. In the adapter approach, further oligonucleotides with the same B–C adapter could be cloned and the number of the cloning cycles is theoretically unlimited. In the polycloning site approach, the number of the cloning cycles is limited to the number of suitable restriction sites present in the vector [B, D and E in Fig. 2(b)].

Applications of the consecutive cloning methods. The examples described here include artificial genes for human insulin A and B chains (adapter method) and for a precursor of the vasoactive intestinal polypeptide, VIP (polycloning site method). The nucleotide sequences and the assembly plan for these genes are shown in Fig. 3.

Human insulin A chain gene. The 5′-terminal part of this gene was obtained by direct cloning of a 45-mer oligonucleotide (IA1) into BamHI and PstI-cleaved M13mp10 vector. The recombinant M13mp10IA1 vector was used to clone IA2 and to complete the gene. IA2 oligonucleotide was ligated through its 3′-terminal PstI extension to the PstI–HindIII adapter. The five-nucleotides-long PstI cohesive end in the adapter was employed to ensure more efficient oligonucleotide – adapter ligation and to avoid self-ligation of the adapter at this end. At the same time, self-ligation of both the adapter and the oligonucleotide–adapter adduct could take place at the phosphorylated HindIII sites. Therefore HindIII cleavage of the reaction mixture was performed before the isolation of the reaction product by non-denaturing polyacrylamide gel electrophoresis (Fig. 4). The isolated partial duplex containing a 5′-terminal single-stranded region (IA2 sequence) and a duplex region bordered by a PstI recognition site and a HindIII cohesive end was then cloned into PstI and HindIII cleaved M13mp10IA1 vector. The M13mp10IA1 vector showed a lac$^+$ phenotype, but insertion of the IA2+adapter sequence shifted the reading-frame of the encoded alpha-peptide. Selection for the recombinants could therefore take advantage of the simple blue-to-white colour reaction.[20] 12 white plaques were checked by nucleotide sequencing and five of them were shown to be the expected M13mp10IA containing the properly joined IA1 and IA2 sequences.

Human insulin B gene. The pUC8IB1 vector containing a 5′-terminal part of the insulin B gene was cleaved with PstI

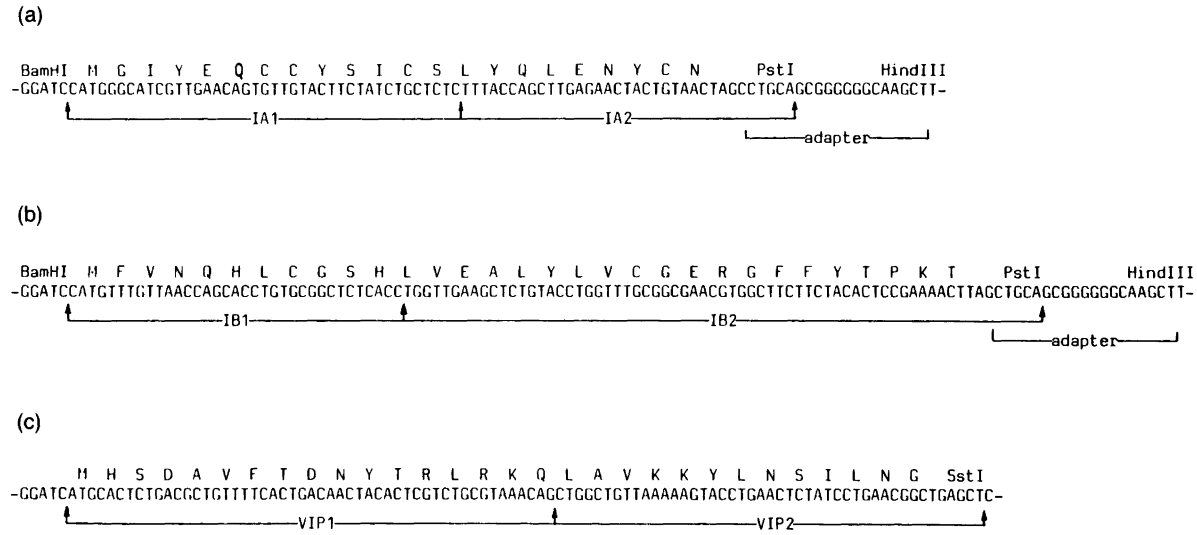

Fig. 3. Sequence and assembly plan of the synthetic genes: (a) M13mp10IA clone containing human insulin A chain gene between the BamHI and PstI sites of M13mp10; (b) pUC8IB clone (pUC8 vector, BamHI and PstI sites); (c) pPEXVIP clone (pPEX vector, BamHI and SstI sites). The arrows show the division of the synthetic genes to the corresponding single-stranded oligonucleotides.

and HindIII. The 67-mer IB2 oligonucleotide ligated with the PstI-HindIII adapter was cloned into this linear vector as described above. Recombinants were selected by colony hybridization using the 5′-[^{32}P]-labelled 67-mer probe. Six hybridizing recombinants were analysed by nucleotide sequencing and two of them were found to contain the expected sequence coding for the whole insulin B chain. This recombinant is denoted as pUC8IB.

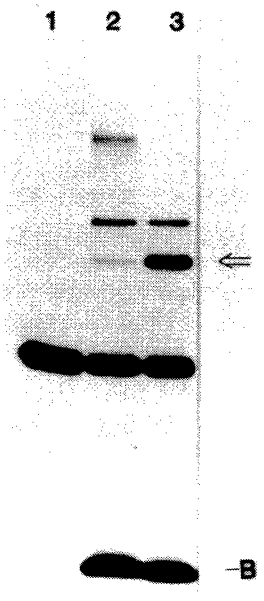

Fig. 4. Ligation of IA2 oligonucleotide with the PstI–HindIII adapter: lane 1, 5′-[^{32}P] labelled IA2; lane 2, reaction mixture obtained after ligation; lane 3, ligated mixture after HindIII cleavage. Electrophoresis was on a 10 % non-denaturing gel followed by autoradiography. B = bromophenol blue marker, which comigrates with the unchanged adapter. The arrow shows the band representing the IA2–adapter partial duplex.

VIP precursor gene. Cloning of the VIP1 oligonucleotide into M13mp18 vector to obtain M13mp18VIP1 has already been described in the first example in this paper. The recombinant vector M13mp18VIP1 could have been used to clone the VIP2 oligonucleotide between the KpnI and SstI sites, but we decided to clone the whole gene directly into the pPEX *E. coli* expression vector.[14] This vector contains an array of unique restriction sites for BamHI, KpnI, SstI, ApaI and EcoRI enzymes. VIP1 oligonucleotide was cloned between the BamHI and KpnI sites of pPEX as described for M13mp18 vector, but the selection was performed by colony hybridization. The resulting pPEXVIP1 recombinant (obtained in 12 %) was then cleaved with KpnI and SstI and the VIP2 oligonucleotide, extended at its 3′-terminus with an AGCT sequence characteristic for SstI site, was ligated with the SstI site of the vector. Klenow polymerase + dNTP reaction converted the VIP2 oligonucleotide into a duplex with concomitant removal of the 3′-protruding extension of the cloned VIP1 oligonucleotide. Blunt-end ligation followed by transformation and selection with the [^{32}P]-labelled VIP2 probe identified positive colonies. Five of them were checked by nucleotide sequencing and in three cases the correct VIP sequence was found. Expression of the pPEXVIP recombinant in JM101 cells followed by positive immunoreaction of the β-galactosidase–VIP fusion protein with antibodies directed against VIP[14] also indicated the presence of the correct VIP gene in the recombinant.

Discussion

Conversion of single-stranded oligonucleotides into cloned duplexes has been performed by a method which includes the following steps: (1) ligation of the oligonucleotide through its 3′-terminal extension to the doubly-cleaved vector; (2) conversion of the single-stranded regions into duplexes; (3) blunt-end ligation; (4) transformation and selection of the recombinants. The cloning procedure is simple and easy to perform as no buffer change or precipitation is included, only addition of further components is required. The overall cloning yield varied between 10 and 25 %, but the proportion of the expected recombinants among the selected clones was 50–75 %.

Compared with other methods when single-strand oligonucleotides are converted into duplexes by either *in vitro*[8] or *in vivo*[9–12] processes, this method requires substantially fewer extra sequences to be synthesized. In our case only the four-nucleotide-long 3′-terminal extension (the central part of a six-base-pair recognition site) can be considered as a redundant synthetic sequence. For a duplex of N length, the percentage of the chemical synthesis saved can be calculated by the formula $100*(N-4)/2N$. Even for $N=40$, which requires a relatively short (44-mer) oligonucleotide by the standards of the present synthetic techniques, as much as 45 % of the chemical work is saved. This figure compares favourably with other methods where much longer oligonucleotides (over 100 bases) are used and the saving obtained is 42 %[10] and 30 %,[12] respectively.

If the cloned duplex is to be isolated by the respective enzymatic cleavages, the 5′- and 3′-terminal (preceding the extension) bases should be defined so that they regenerate the respective recognition sites. In the strategies we use for gene synthesis, the cloned oligonucleotide is not isolated as a duplex but remains part of the vector which in turn is used for the cloning of the next incoming oligonucleotide. Such consecutive applications have been realized in two methods (the adapter and the polycloning site methods), using M13 phage, plasmid and *E. coli* expression vectors. Each oligonucleotide carries a 3′-terminal extension, but this is eliminated when the oligonucleotides are joined in a cloning vector. Only one nucleotide preceding the 3′-terminal, four-nucleotide-long extension remains in the cloned sequence. This nucleotide should be defined so that it regenerates the restriction site used previously in the oligonucleotide cloning, permitting the use of this regenerated site in the next step of the cyclic application. For example, when oligonucleotides with PstI-like CTGCA 3′-terminal extensions are sequentially cloned, the defined nucleotide which remains at the joining point is C.

Recently, an alternative approach for gene synthesis based on self-priming oligonucleotides has been described,[30] which is similar to the adapter method presented here. Our approach shares most of the advantages detailed in Ref. 30, but it is different in that that the very same adapter molecule can be joined covalently to many oligonucleotides, while different self-priming stem-loop structures have to be synthesized as integral parts of each oligonucleotide.[30] Moreover, the fill-in reaction in our case takes place so that the oligonucleotide is already connected to the cloning vector, therefore the use of different restriction enzyme pairs for each oligonucleotide, the isolation of the cloned parts and their assembly by other methods[2,31] are eliminated.

In conclusion, besides the saving (nearly 50 %) in the synthetic work, the gene assembly can also be simplified when single-stranded oligonucleotides are sequentially cloned according to the presented methods. So far at least 50 oligonucleotides, 30 to 100 nucleotides long have successfully been cloned in our laboratories using the techniques described here.

Acknowledgements. Thanks are due to Gy. Hollo Mozes, M. Rozsavölgyi and E. Novak for their skilful technical assistance. Part of this work was supported by a grant (OTKA552) from the Hungarian Academy of Sciences to A. S. Travel grants and a fellowship to M.-L. T. from Stockholm University and from the Swedish Institute are acknowledged.

References

1. Engels, J. W. and Uhlmann, E. *Angew. Chem., Int. Ed. Engl. 28* (1989) 716.
2. Khorana, H. G. *Science 203* (1979) 614.
3. Ferretti, L., Karnik, S. S., Khorana, H. G., Nassal, M. and Oprian, D. D. *Proc. Natl. Acad. Sci. USA 83* (1986) 599.
4. Bell, L. D., Smith, J. C., Derbyshire, R., Finlay, M., Johnson, I., Gilbert, R., Slocombe, P., Cook, E., Richards, H., Clissold, P., Meredith, D., Powell-Jones, C. H., Dawson, K. M., Carter, B. L. and McCullagh, K. G. *Gene 63* (1988) 155.
5. Brousseau, R., Scarpulla, R., Sung, W., Hsiung, H. M., Narang, S. A. and Wu, R. *Gene 17* (1982) 279.
6. Rossi, J. J., Kierzek, R., Huang, T., Walker, P. A. and Itakura, K. *J. Biol. Chem. 257* (1982) 9226.
7. Rink, H., Liersch, M., Sieber, P. and Meyer, F. *Nucleic Acids Res. 12* (1984) 6369.
8. Oliphant, A. R., Nussbaum, A. L. and Struhl, K. *Gene 44* (1986) 177.
9. Derbyshire, K. M., Salvo, J. J. and Grindley, N. D. F. *Gene 46* (1986) 146.
10. Adams, S. E., Johnson, I. D., Braddock, M., Kingsman, A. J., Kingsman, S. M. and Edwards, R. M. *Nucleic Acids Res. 16* (1988) 4287.
11. Weiner, M. P. and Scheraga, H. A. *Nucleic Acids Res. 17* (1989) 7113.
12. Mandecki, W. and Bolling, T. J. *Gene 68* (1988) 101.
13. Simoncsits, A., Kalman, M., Kari, C. and Cserpan, I. *Chem. Abstr. 105* (1986) 222196u and patent applications referred to therein.
14. Simoncsits, A., Tjörnhammar, M.-L., Kalman, M., Cserpan, I., Gafvelin, G. and Bartfai, T. *Eur. J. Biochem. 178* (1988) 343.
15. Kalman, M., Cserpan, I., Bajszar, G., Dobi, A., Horvath, E., Pazman, C. and Simoncsits, A. *Nucleic Acids Res. 18* (1990) 6075.
16. Simoncsits, A. In: Townsend, L. B. and Tipson, R. S., Eds., *Synthetic Procedures in Nucleic Acid Chemistry,* Wiley, Chichester, UK, Vol. 4. *In press.*

17. Norrander, J., Kempe, T. and Messing, J. *Gene 26* (1983) 101.
18. Yanisch-Perron, C., Vieira, J. and Messing, J. *Gene 33* (1985) 103.
19. Vieira, J. and Messing, J. *Gene 19* (1982) 259.
20. Gronenborn, B. and Messing, J. *Nature (London) 272* (1978) 375.
21. Hanahan, D. J. *Mol. Biol. 166* (1983) 557.
22. Maniatis, T., Fritsch, E. F. and Sambrook, J. *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 1982.
23. Sproat, B. S. and Bannwarth, W. *Tetrahedron Lett. 24* (1983) 5771.
24. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. and Roe, B. A. *J. Mol. Biol. 143* (1980) 161.
25. Biggin, M. D., Gibson, T. J. and Hong, G. F. *Proc. Natl. Acad. Sci. USA 80* (1983) 3963.
26. Hong, G. F. *Bioscience Rep. 2* (1982) 907.
27. Chen, Y. E. and Seeburg, P. H. Z. *DNA 4* (1985) 165.
28. Boros, I., Lukacsovich, T., Baliko, G. and Venetianer, P. *Gene 42* (1986) 97.
29. Kalnius, A., Otto, K., Ruther, U. and Müller-Hill, B. *EMBO J. 4* (1983) 593.
30. Uhlmann, E. *Gene 71* (1988) 29.
31. Hostomosky, Z., Smrt, J., Arnold, L., Tocik, Z. and Paces, W. *Nucleic Acids Res. 15* (1987) 4849.

Received May 17, 1990.