Nucleic Acids Research, 1998, Vol. 26, No. 5 **1223–1227**

Codon-based mutagenesis using dimer-phosphoramidites

Philippe Neuner*, Riccardo Cortese[1] and Paolo Monaci[1]

Chemistry Department and [1]Biotechnology Department, Istituto di Ricerche di Biologia Molecolare P. Angeletti, Via Pontina km 30.600, 00040 Pomezia (Roma), Italy

Received November 17, 1997; Revised and Accepted January 16, 1998

ABSTRACT

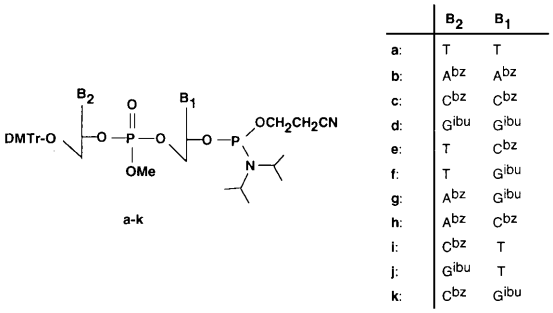
A new approach for the synthesis of randomized DNA sequences containing the 20 codons corresponding to all natural amino acids is described. The strategy is based on the use of dinucleotide phosphoramidite building blocks within a resin-splitting procedure. Through this protocol, a minimal number of seven dimers is sufficient to encode all 20 natural amino acids. This synthesis procedure is extremely flexible and allows codon usage from different hosts to be accommodated.

INTRODUCTION

Oligonucleotide-directed mutagenesis is the method most commonly used to generate mutants of a recombinant protein. Several strategies have been developed to introduce defined or random residues at given positions of a polypeptide coding sequence. In standard oligonucleotide-directed mutagenesis schemes, a randomized DNA sequence is synthesized by sequentially coupling a mixture of the four nucleoside precursors to the growing oligonucleotide. In this way all 64 possible codon sequences are generated (NNN, including 41 redundant and 3 stop codons). This strategy can be improved by exploiting the third position redundancy of many codon assignments. By using all four nucleosides in the first two codon positions, but only G and C or A and T in the third position (NNG/C or NNA/T), the resulting mixture contains 32 triplets encoding all 20 amino acids, with 11 redundant and 1 stop codons. In this manner, the bias in favour of the amino acids encoded by multiple codon sequences is maintained, and the presence of a stop codon will produce truncated amino acid sequences upon translation. This event, which occurs with a frequency of (n/32) where n is the number of amino acids of the randomized sequence, considerably limits the complexity that can be achieved for long randomized peptide libraries. With this strategy, introducing subsets of the 20 amino acids at a given position in the molecule, e.g. to exclude the codon corresponding to the wild type sequence (1), is limited to only those combinations that can be generated through the synthesis of mixtures of monomers. An alternative approach is based on the synthesis of individual codon sequences to the growing oligonucleotide on separate columns, as described for the synthesis of random peptide libraries (2). After synthesis of each codon, the beads from all the columns are mixed together, split again and then repacked into new columns to synthesize the next codon (3–5). Such codon-based mutagenesis eliminates redundant codons, thus creating more compact repertoires of mutants, since only $(20/32)^n$ of the clones are required to include all possible residues for an n-long amino acid region. In addition, the absence of stop codons prevents premature termination of the peptide sequences generated, thus removing limitations to the complexity of the library. This resin-splitting method allows randomization of any codon position to be modulated by varying the proportions of starting materials and/or reaction products mixed at each step. However, the benefits are counteracted by having to use several columns and a procedure that becomes increasingly labourious with the complexity of the mutagenesis scheme, making the task practically impossible.

A third group of methods is based on the use of 20 pre-synthesized codons as monomeric units. Methods have been published describing the preparation of trinucleotide phosphoramidites and their use in synthesizing randomized oligonucleotide sequences by automated DNA synthesis methodology (6–10). This would appear an obvious route to easily achieve fine control over the presence of any set of residues at any given position, but the synthesis and efficient coupling of trinucleotide blocks is not a straightforward process. Early attempts to use triplets for the generation of protein mutants reported low coupling efficiency, as well as deletions in the final product (6). Lyttle and co-workers (7) reported the synthesis of five cyanoethyl phosphoramidite codon triplets and their use for oligonucleotide synthesis. Their protocol showed a modest and uneven coupling efficiency of the triplets and was found to generate a significant amount of single-base insertions. Ono *et al.* (8) designed the synthesis of the antisense codon triplets in the 5′–3′ direction. These anti-codons were then converted into the sense strand by *in vitro* replication methods. In this case too, more study would be required to establish optimal conditions for coupling reactions to achieve equimolar incorporation of the codons. Virnekäs *et al.* (9) described the synthesis of 20 pre-made trinucleotide phosphoramidites and their use in conventional 3′–5′ solid-phase DNA synthesis. By using slightly longer reaction times and performing double coupling cycles, good yields of randomized DNA oligonucleotides were obtained with conventional automated solid-phase synthesis. However, synthesis of the 20 triplet blocks is long and complicated

*To whom correspondence should be addressed. Tel: +39 6 91 093 335; Fax: +39 6 91 093 225; Email: neuner@irbm.it



	B_2	B_1
a:	T	T
b:	A^{bz}	A^{bz}
c:	C^{bz}	C^{bz}
d:	G^{ibu}	G^{ibu}
e:	T	C^{bz}
f:	T	G^{ibu}
g:	A^{bz}	G^{ibu}
h:	A^{bz}	C^{bz}
i:	C^{bz}	T
j:	G^{ibu}	T
k:	C^{bz}	G^{ibu}

Figure 1. Chemical structure of the 11 dinucleotide phosphoramidite building blocks.

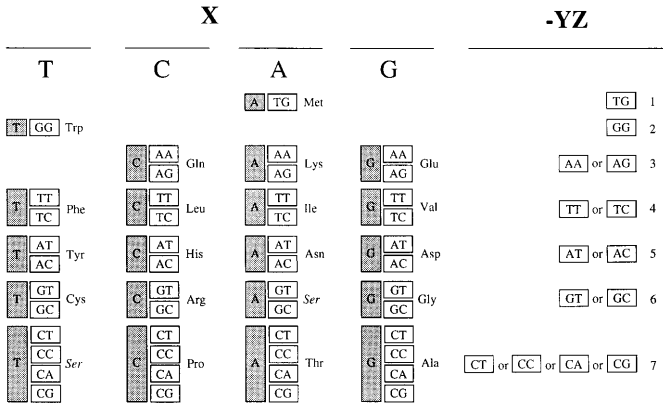


Figure 2. Sets of -YZ dinucleotides whose combination with a nucleoside X at the first position in the codon sequence encode the 20 natural amino acids. Ser codon (indicated in italics) can be generated by -CG or -GT dinucleotide units.

and these triplets are in any case obtained in low yields. Recently, a method for the large scale synthesis of trinucleotide phosphotriester synthons was reported (10). This protocol was proved efficient by generating a library of single chain Fv fragments in which the eight amino acid CDR3 region of the heavy chain was randomized (11).

Here we describe a novel codon-based mutagenesis strategy using dinucleotide phosphoramidite building blocks within a resin-splitting framework. We show that only seven dinucleotide building blocks are required to encode all the 20 natural amino acids. We discuss the advantages and extreme flexibility of this approach and describe the application of this procedure to the synthesis of a randomized sequence in *Escherichia coli*.

MATERIALS AND METHODS

General materials and methods

The synthesis of the dinucleotide building blocks was essentially accomplished as described by Kumar *et al*. (12), and the chemical structure of the 11 dimers (a–k) is reported in Figure 1. All dimers had the expected ^1H and ^{31}P spectroscopic properties.

Automated oligonucleotide synthesis was performed on an Applied Biosystems 394 DNA/RNA synthesizer, using the standard 40 nmol scale synthesis protocols, with coupling time of the dinucleotide units extended to 2 min. The synthesis reagents were obtained from Applied Biosystems. Mixtures of dinucleotide phosphoramidites were dried for 15 h over phosphorous pentoxide under vacuum, dissolved in acetonitrile (<10 p.p.m. of water, Labscan) and attached to positions 5, 6, 7 and 8 of the DNA synthesizer. The oligonucleotides synthesized using the dinucleotide building blocks were treated with thiophenol (thiophenol/dioxane/triethylamine) to cleave the methyl phosphate ester protecting group, prior to ammonia treatment (16 h at 55°C). Volatile components were evaporated and the crude oligonucleotide was resuspended in H_2O. Oligonucleotides were used without further purification.

Recombinant DNA methods

Klenow polymerase and T4 DNA ligase were purchased from New England Biolabs and used as recommended by the manufacturer. dNTPs were obtained from Boehringer Mannheim. Methods for plasmid purification, enzymatic reactions, cloning and bacterial transformation were performed as described (13).

Oligonucleotides A (5′-AGTCGCGP$_2$P$_1$TCGACCT-3′ where P_1 and P_2 indicate a mixture of the 20 codons most abundant in highly expressed genes of *E.coli* and B (5′-CGCGACTAGGTCGA-3′) were enzymatically phosphorylated at their 5′ end, mixed in equimolar amounts, denatured at 95°C and then allowed to gradually cool at 15°C for annealing. The oligonucleotide mixture was then subjected to enzymatic ligation and to a filling-in reaction with Klenow polymerase. The blunt-end fragments thus produced were cloned in the *Eco*RV site of phagemid pBSks+ (14). The ligation mix was enriched for recombinant clones through *Eco*RV digestion prior to transformation in XL1-blue competent bacterial cells (15). Recombinant clones were identified by colour screening on Xgal/IPTG/ampicillin plates. PCR amplification of the cloned sequence and gel electrophoresis analysis of the PCR products identified inserts composed of multiple copies of the oligonucleotide A. DNA sequencing was performed with dydeoxy terminator Taq cycle sequencing kit (Perkin Elmer) on an Applied Biosystems 373 automated DNA sequencer.

RESULTS AND DISCUSSION

Rationale

A triplet codon sequence (XYZ) can be formed by the association of one nucleoside at 5′ end (X) and one dinucleotide at 3′ end (-YZ). According to the genetic code, a minimal number of seven -YZ dimers associated with the proper nucleoside X at the first position in the codon sequence is sufficient to encode the 20 natural amino acids. It is possible to identify 65 sets of seven different -YZ dimers that share this property, obtained from the combination of one dimer from each of the seven -YZ groups listed in Figure 2.

For example, the 20 codons can be generated by using the seven dimer blocks -TG, -GG, -AG, -TC, -AC, -GC and -CG associated with the proper 5′ nucleoside, but also by the seven dimer blocks -TG, -GG, -AA, -TC, -AC, -GT and -CG. Any of these 65 combinations can be used according to different criteria. For example, the combination of the seven -YZ dimer blocks (-TC, -TG, -CC, -AC, -AG, -GC and -GG) with the proper X nucleoside at the 5′, generates the triplets most abundant in eukaryotic

Table 1. Codon usage in eukaryotes (**A**), *E.coli* (**B**) and yeast (**C**)

A

X	T	C	A	G	Y	Z
T	20 TTT Phe	5 CTT Leu	18 ATT Ile	7 GTT Val	T	T
	80 TTC Phe	26 CTC Leu	**77 ATC Ile**	25 GTC Val	T	C
	2 TTA Leu	3 CTA Leu	5 ATA Ile	5 GTA Val	T	A
	6 TTG Leu	**58 CTG Leu**	**100 ATG Met**	**64 GTG Val**	T	G
C	13 TCT Ser	19 CCT Pro	14 ACT Thr	17 GCT Ala	C	T
	28 TCC Ser	**48 CCC Pro**	**57 ACC Thr**	**53 GCC Ala**	C	C
	5 TCA Ser	16 CCA Pro	14 ACA Thr	13 GCA Ala	C	A
	9 TCG Ser	17 CCG Pro	15 ACG Thr	17 GCG Ala	C	G
A	26 TAT Tyr	21 CAT His	22 AAT Asn	25 GAT Asp	A	T
	74 TAC Tyr	**79 CAC His**	**78 AAC Asn**	**75 GAC Asp**	A	C
	23 TAA STOP	12 CAA Gln	18 AAA Lys	25 GAA Glu	A	A
	21 TAG STOP	**88 CAG Gln**	**82 AAG Lys**	**75 GAG Glu**	A	G
G	32 TGT Cys	7 CGT Arg	10 AGT Ser	12 GGT Gly	G	T
	68 TGC Cys	**37 CGC Arg**	**34 AGC Ser**	**50 GGC Gly**	G	C
	55 TGA STOP	6 CGA Arg	10 AGA Arg	14 GGA Gly	G	A
	100 TGG Trp	21 CGG Arg	18 AGG Arg	24 GGG Gly	G	G

B

X	T	C	A	G	Y	Z
T	24 TTT Phe	4 CTT Leu	17 ATT Ile	**51 GTT Val**	T	T
	76 TTC Phe	7 CTC Leu	**83 ATC Ile**	7 GTC Val	T	C
	2 TTA Leu	0 CTA Leu	0 ATA Ile	26 GTA Val	T	A
	3 TTG Leu	**83 CTG Leu**	**100 ATG Met**	16 GTG Val	T	G
C	34 TCT Ser	8 CCT Pro	35 ACT Thr	**35 GCT Ala**	C	T
	37 TCC Ser	0 CCC Pro	**55 ACC Thr**	10 GCC Ala	C	C
	2 TCA Ser	15 CCA Pro	4 ACA Thr	28 GCA Ala	C	A
	3 TCG Ser	**77 CCG Pro**	7 ACG Thr	26 GCG Ala	C	G
A	25 TAT Tyr	17 CAT His	6 AAT Asn	33 GAT Asp	A	T
	75 TAC Tyr	**83 CAC His**	**94 AAC Asn**	**67 GAC Asp**	A	C
	0 TAA STOP	14 CAA Gln	**74 AAA Lys**	**78 GAA Glu**	A	A
	0 TAG STOP	**86 CAG Gln**	26 AAG Lys	22 GAG Glu	A	G
G	49 TGT Cys	**74 CGT Arg**	3 AGT Ser	**59 GGT Gly**	G	T
	51 TGC Cys	25 CGC Arg	20 AGC Ser	38 GGC Gly	G	C
	0 TGA STOP	1 CGA Arg	0 AGA Arg	0 GGA Gly	G	A
	100 TGG Trp	0 CGG Arg	0 AGG Arg	2 GGG Gly	G	G

C

X	T	C	A	G	Y	Z
T	27 TTT Phe	3 CTT Leu	**52 ATT Ile**	**56 GTT Val**	T	T
	73 TTC Phe	0 CTC Leu	46 ATC Ile	37 GTC Val	T	C
	18 TTA Leu	7 CTA Leu	2 ATA Ile	3 GTA Val	T	A
	69 TTG Leu	2 CTG Leu	**100 ATG Met**	4 GTG Val	T	G
C	**52 TCT Ser**	18 CCT Pro	**50 ACT Thr**	**65 GCT Ala**	C	T
	33 TCC Ser	2 CCC Pro	43 ACC Thr	28 GCC Ala	C	C
	6 TCA Ser	**80 CCA Pro**	6 ACA Thr	6 GCA Ala	C	A
	1 TCG Ser	1 CCG Pro	1 ACG Thr	1 GCG Ala	C	G
A	19 TAT Tyr	35 CAT His	22 AAT Asn	**52 GAT Asp**	A	T
	81 TAC Tyr	**65 CAC His**	**78 AAC Asn**	48 GAC Asp	A	C
	80 TAA STOP	**95 CAA Gln**	22 AAA Lys	**90 GAA Glu**	A	A
	10 TAG STOP	5 CAG Gln	**78 AAG Lys**	10 GAG Glu	A	G
G	**89 TGT Cys**	15 CGT Arg	5 AGT Ser	**91 GGT Gly**	G	T
	11 TGC Cys	0 CGC Arg	4 AGC Ser	6 GGC Gly	G	C
	9 TGA STOP	0 CGA Arg	**83 AGA Arg**	2 GGA Gly	G	A
	100 TGG Trp	0 CGG Arg	1 AGG Arg	1 GGG Gly	G	G

X, Y and Z indicate the nucleosides in the first, second and third positions in the triplet, respectively. Most abundant codons in highly expressed genes are indicated in bold. Triplets derived from the combination of -YZ dimer blocks with the proper X nucleoside at the 5′ end are highlighted.

highly-expressed genes, as described in Table 1A. The formation of a triplet through the X–YZ system offers considerable flexibility and allows the codons most frequent in highly expressed genes of different hosts to be accommodated. The 11 -YZ dinucleotide blocks -TT, -TC, -TG, -CT, -CC, -CG, -AC, -AA, -AG, -GT and -GG associated with the proper X nucleoside at the 5′, form the 20 codons most abundant in highly expressed genes of *E.coli*, as illustrated in Table 1B.

Similarly, by combining the 11 -YZ dinucleotide blocks (-TT, -TC, -TG, -CT, -CA, -AC, -AA, -AG, -GT, -GA and -GG) with a proper nucleoside X at the first codon position, the 20 codons most frequent in highly expressed genes in yeast are generated, as indicated in Table 1C.

It is worth noting that the association of a 5′ dinucleotide unit (XY-) with a 3′ nucleoside (Z) to form a XY–Z triplet requires a minimum of 13 different dimer units to represent all 20 amino acids.

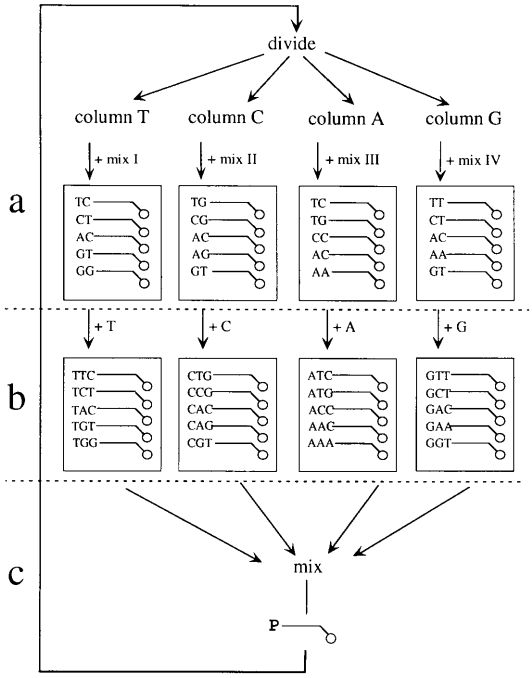


Figure 3. Strategy for the synthesis of a randomized DNA sequence containing the 20 most abundant codons in highly expressed genes of *E.coli*.

Codon-based mutagenesis strategy utilizing dinucleotide blocks

We wanted to exploit the above described properties of the X–YZ method of codon formation to establish a general protocol for codon-based mutagenesis. To this end, we opted for a scheme of codon synthesis based on the association between a limited number of pre-formed 3′ dimeric units and 5′ nucleosides, and an X–YZ strategy of codon formation combined with a resin-splitting method, thus drastically simplifying the procedures and diverse steps required to construct a randomized DNA sequence. The method we propose allows any set of the 20 amino acid coding triplets to be introduced at any given position during synthesis of the oligonucleotide chain. Figure 3 outlines the strategy for the synthesis of a randomized DNA sequence containing the 20 codons most abundant in highly expressed genes in *E.coli*. The 3′-end flanking region of the oligonucleotide is synthesized on an automated DNA synthesizer using conventional phosphoramidite chemistry on solid phase support. When the randomized sequence is synthesized, an equivalent amount of the resin is loaded on four columns (columns T, C, A and G, respectively). Four pre-selected mixtures of dinucleotide phosphoramidite building blocks are composed according to the combinations highlighted in Table 1B and separately coupled to their respective synthesis resin (Fig. 3a). Then, the mononucleoside phosphoramidite corresponding to the first base of the codon is coupled to its respective synthesis resin (Fig. 3b). The phosphoramidite nucleoside T is coupled to the growing oligonucleotides with the -TC, -CT, -AC, -GT or -GG sequence at their 5′ end (column T), to generate the five codons TTC (Phe), TCT (Ser), TAC (Tyr), TGT (Cys) and TGG (Trp), respectively. Similarly, the phosphoramidite nucleoside C is

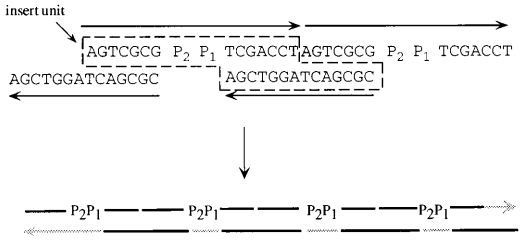


Figure 4. Scheme of construction of the 'head-to-tail' polymer of the insert unit.

coupled to the growing oligonucleotides with the -TG, -CG, -AC, -AG and -GT sequence at their 5′ end (column C), to generate the five codons CTG (Leu), CCG (Pro), CAC (His), CAG (Gln) and CGT (Arg), respectively. The phosphoramidite nucleoside A is coupled to the growing oligonucleotides with the -TC, -TG, -CC, -AC and -AA sequence at their 5′ end (column A), to generate the five codons ATC (Ile), ATG (Met), ACC (Thr), AAC (Asn) and AAA (Lys), respectively. Finally, the phosphoramidite nucleoside G is coupled to the growing oligonucleotides with the -TT, -CT, -AC, -AA and -GT sequence at their 5′ end (column G), to generate the five codons GTT (Val), GCT (Ala), GAC (Asp), GAA (Glu) and GGT (Gly), respectively. The resin of the four columns is then pooled, mixed and four new columns are loaded with an equivalent amount of the resin mixture (Fig. 3c). This procedure is repeated until the desired length of the randomized sequence has been reached. Finally, the four synthesis resins are pooled, the 5′-end flanking region is synthesized, the oligomer deprotected and cleaved from the support. The same method can be adapted to any specific requirements by properly choosing and distributing the -YZ dimeric blocks. For example, codons for acidic amino acids (Asp and Glu) can be easily excluded by omitting the -AC and -AA dimer blocks from mix IV in the scheme of synthesis described in Figure 3. This protocol of codon-based mutagenesis has clear advantages over the methods based on the synthesis of trinucleotide synthons. The dinucleotide building blocks are usually obtained in better yields than trinucleotide units, and pre-made dimers show a high coupling efficiency when incorporated in oligonucleotide synthesis. At the same time, the use of these dimer units simplifies the resin-splitting method of mutagenesis using mononucleoside phosphoramidite (4). In fact, the number of synthesis columns decreases from 10 to 4, and the corresponding splitting and mixing operations are also reduced, thus improving the overall yield and the quality of the product.

To test the feasibility of this mutagenesis method, we synthesized an oligonucleotide (oligonucleotide A) containing two random amino acid positions (P_2 and P_1; Fig. 4). On the basis of the *E.coli* codon usage, the 11 dimers highlighted in Table 1B were chosen as synthons.

Preparation of the dinucleotide phosphoramidite mixtures

The content of dinucleotide synthon mixtures I–IV was as follows: mixture I contained the five dimer-amidites -TC, -CT, -AC, -GT and -GG; mixture II the five dimer-amidites -TG, -CG, -AC, -AG and -GT; mixture III the five dimer-amidites -TC, -TG, -CC, -AC; and -AA and mixture IV the five dimer-amidites -TT, -CT, -AC, -AA and -GT. As already reported by other authors for the coupling of trinucleotide synthons (8–10), the activated amidite species differ in

Table 2. Composition and concentration of the dimer–amidite mixtures

mix I		mix II		mix III		mix IV	
-TC	0.0225 M	-TG	0.0150 M	-TC	0.0225 M	-TT	0.0225 M
-CT	0.0150 M	-CG	0.0150 M	-TG	0.0150 M	-CT	0.0150 M
-AC	0.0150 M	-AC	0.0150 M	-CC	0.0150 M	-AC	0.0150 M
-GT	0.0150 M	-AG	0.0225 M	-AC	0.0150 M	-AA	0.0150 M
-GG	0.0300 M	-GT	0.0150 M	-AA	0.0150 M	-GT	0.0150 M
conc. = 0.0975 M		conc. = 0.0825 M		conc. = 0.0825 M		conc. = 0.0825 M	

Table 3. Frequencies of the 20 codons in the synthesized random DNA sequence

								f/n	-YZ
						GTT	14	14.0	-TT
TTC	6			ATC	11			8.5	-TC
		CTG	12	ATG	7			9.5	-TG
TCT	10					GCT	5	7.5	-CT
				ACC	7			7.0	-CC
		CCG	5					5.0	-CG
TAC	10	CAC	11	AAC	7	GAC	11	9.8	-AC
				AAA	11	GAA	11	11.0	-AA
		CAG	13					13.0	-AG
TGT	3	CGT	7			GGT	3	4.3	-GT
TGG	6							6.0	-GG

Average frequencies (f/n) for the triplets derived from the same dimer unit (-YZ) are reported in the rightmost columns.

their coupling reactivity to the growing oligomer. Concentrations of the dimer building blocks (Table 2) were empirically determined to achieve an equal representation of all the 20 codons in the randomized DNA sequence. The dimer building block units were carefully dried, dissolved in acetonitrile (at a concentration of 0.0975 M for mixture I and 0.0825 M for mixtures II, III and IV).

Synthesis of the oligonucleotide with randomized sequence

The 3′-end flanking region of the oligonucleotide A (TCGACCT-3′) was synthesized in parallel on four columns (40 nmol scale) on a DNA synthesizer using conventional phosphoramidite chemistry on solid phase support. The first randomized position P_1 was synthesized as described above, according to the scheme reported in Figure 3. The cycle was then repeated to form P_2 in the growing oligomer sequence. The 5′-end flanking region (5′-AGTCGCG) of the synthetic oligonucleotide was subsequently synthesized on all four columns, the synthesis resins were pooled, the oligomer deprotected and cleaved from the support, as described in Materials and Methods.

Sequence analysis of the randomized DNA sequence

In order to analyze a sufficient number of codons, we synthesized the 14mer oligonucleotide B (5′-CGCGACTAGGTCGA-3′) using standard phosphoramidite chemistry. The structure of oligonucleotide B was designed so that its 3′- and 5′-halves (each 7 nucleotides long) were complementary to the 3′- and the 5′-constant regions of oligonucleotide A (Fig. 4). This feature allowed us to generate a 'head-to-tail' double stranded polymer by annealing and extending with DNA polymerase an equimolar mixture of the two oligonucleotides. The polymerization product was cloned into a phagemid vector and 20 selected recombinants from this mini-library were sequenced according to the procedure described in Materials and Methods. By analyzing clones containing multiple copies of the randomized sequence P_2P_1, we

were able to identify a sufficient number of codons performing a limited number of sequencing reactions. In Table 3 we report the frequency of the 20 triplets observed in the 170 codons analyzed, where all the expected 20 codon sequences were detected. We also calculated the average frequency of triplets derived from coupling an X nucleoside with the same -YZ dimer unit. These data show that the 11 dimers are incorporated with comparable frequency, although some of them appear to be either under- (e.g. -GT) or over-represented (e.g. -TT). These minor variations can be corrected by empirically modifying the proportion of dimer building blocks present in the original mixtures.

In conclusion, we have developed a versatile route to prepare randomized synthetic DNA sequences using dinucleotide building blocks. This method has been successfully applied to the construction of DNA libraries with three to five randomized codon positions that are currently being screened in the laboratory. We believe that the strategy proposed here could be extremely effective in creating molecular diversity by mutagenesis, and particularly viable where a considerable diversity of sequences is desired, such as the construction of random peptide libraries displayed on phage, randomization of protein segments or affinity maturation of ligands.

ACKNOWLEDGEMENTS

We wish to thank A. Tramontano, M. Sollazzo, A. Nicosia, V. Matassa and C. Toniatti for their critical reading of the manuscript, G. Galfré for useful discussion and J. Clench for proofreading. We are grateful to M. Nuzzo for technical help and to M. Pezzanera for sequencing analysis.

REFERENCES

1 Huang, W. and Santi, D. V. (1994) *Anal. Biochem*., **218**, 454–457.
2 Lam, K. S., Salmon, S. E., Hersh, E. N., Hruby, V. J., Katzmierski, W. M. and Knapp, R. J. (1991) *Nature*, **354**, 82–84.
3 Glaser, S. M., Yelton D. E. and Huse W. D. (1992) *J. Immunol*., **149**, 3903–3913.
4 Huse, W. D. (1993) *US patent* 5, **264**, 563.
5 Cormack, B. P. and Struhl, K. (1993) *Science*, **262**, 244–248.
6 Sondek, J. and Shortle D. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 3581–3585.
7 Lyttle, M. H., Napolitano E. W., Calio B. L. and Kauvar L. M. (1995) *Biotechniques*, **19**, 274–280.
8 Ono, A., Matsuda, A., Zhao J. and Santi D. (1995) *Nucleic Acids Res*., **23**, 4677–4682.
9 Virnekäs, B., Ge, L., Plückthun, A., Schneider, K. C., Wellnhofer, G. and Moroney, S. E. (1994) *Nucleic Acids Res*., **22**, 5600–5607.
10 Kayushin, A. L., Korosteleva, M. D., Miroshnikov, A. I., Kosch, W., Zubov, D. and Piel, N. (1996) *Nucleic Acids Res*., **24**, 3748–3755.
11 Braunagel, M. and Little, M. (1997) *Nucleic Acids Res*., **25**, 4690–4691.
12 Kumar, G. and Poonian, M. S. (1984) *J. Org. Chem*., **49**, 4905–4912.
13 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14 Short, J. M., Fernandez, J. M., Sorge, J. A. and Huse, W. D. (1988) *Nucleic Acids Res*., **16**, 7583–7600.
15 Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987) *Biotechniques*, **5**, 376.