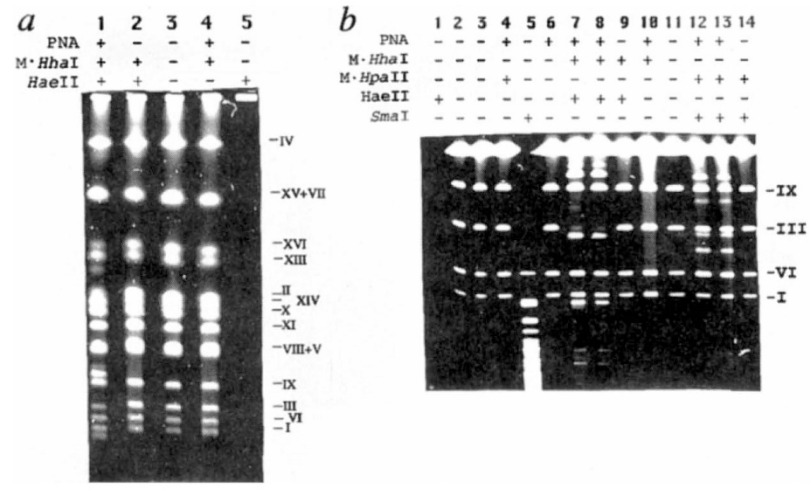
PNA as a rare genome-cutter

SIR—There is a continuing need for techniques that will allow genomic DNA strands to be cut into pieces of between several hundred kilobase pairs (kbp) and a few megabase pairs (Mbp). Naturally occurring restriction enzymes and endonucleases, which usually recognize between four and six contiguous nucleotides, are too nonspecific to yield such large fragments. Here we demonstrate that a peptide nucleic acid (PNA) 'clamp', consisting of two predesigned 8-bp sequences of PNA linked together, can bind strongly and sequence-specifically to DNA, protecting the binding region from methylation. After removal of this bis-PNA, methylated DNA can be cut quantitatively by restriction enzymes at the site selected by the PNA clamp.

PNA, an oligonucleotide mimic in which the common DNA bases are attached by linkers to an *N*-(2-aminoethyl)glycine backbone[1], forms exceptionally stable complexes with duplex DNA[2]. When all PNA bases are pyrimidines, binding to duplex DNA involves the formation of a PNA_2/DNA complex and the displacement of the complementary strand. The complex can be stabilized further by replacing the cytosines in half of the bis-PNA with pseudoisocytosines[3]. The specificity and binding strength of bis-PNA can be used to design rare cutters of unknown genome systems. This method is radically different from existing approaches, which permit DNA cutting only at predetermined sites of 20 bp or more[4–6]. We estimate that our optimum length for the binding sequence is around 8–10 bp.

We tested our procedure using a bis-PNA molecule, H-Lys_2-TJJJJTTJ-$(eg1)_3$-CTTCCCCT-Lys-NH_2 (where J refers to pseudoisocytosine, and eg1 is 8-amino-3,6-dioxaoctanoic acid), which has a unique 8-bp binding site in λ-DNA overlapping one of the 121 naturally occurring 4-bp *Taq*I sites. The bis-PNA protected this *Taq*I site against *Cvi*BIII methylase (which shows the site specificity of *Taq*I methylase), as shown by subsequent cleavage of the λ-DNA into two fragments by *Taq*I (data not shown). It is an indication of the specificity of the cutting process that we found no extra bands in our gels, even though binding of the bis-PNA with only one mismatch would have overlapped another of the *Taq*I-binding sites.

To assess the use of bis-PNA in cutting genomic DNA, we investigated cleavage of the 16 chromosomes of the yeast genome using two combinations of methylation and restriction enzymes: M·*Hha*I/*Hae*II and M·*Hpa*II/*Sma*I (see figure). When bis-PNA and the *Hha*I/*Hae*II pair are used, several strong bands appear (*a* in the figure, lane 1): below chromosome XIII, above chromosome IX and below chromosome I. Chromosomes XVI and II are quantitatively cut. Higher resolution shows that chromosomes I, VI and IX are uncut whereas chromosome III is completely cut (*b* in the figure, lanes 7 and 8).



Application of the proposed technique to the yeast genome. *a*, Results of pulsed-field gel electrophoresis (PFGE) separation of yeast genome (strain PSY316) after applying bis-PNA and the M·*Hha*I/*Hae*II pair of enzymes. Yeast DNA in a low-melting agarose plug was equilibrated with buffer (20 mM MES, 2 mM EDTA, 10 mM NaCl, pH 6.3, 37 °C) and incubated with 0.2 μM bis-PNA (10 h, 37 °C). Samples were equilibrated in buffer containing 25 mM sodium citrate and 5 mM β-mercaptoethanol (pH 7.0, 37 °C), and incubated overnight with 15 U *Hha*I methylase (4 °C, 100 mg ml^{-1} BSA). *S*-adenosylmethionine (300 mM) was added and methylation performed (37 °C, 2.5 h). The reaction was stopped and the PNA–DNA complex was dissociated (50 min at 58 °C in 50 mM Tris–HCl, 500 mM NaCl, 10 mM EDTA pH 8.9 and 1% SDS). Samples were equilibrated in NEBuffer 4, digested with 4 U *Hae*II restriction endonuclease (100 mg ml^{-1} BSA, 4 h, 37 °C), and the reaction was then stopped (20 min, 50 mM EDTA, pH 8.7). Samples were equilibrated in TE buffer, pH 7.5, and loaded onto a 1% agarose gel. The PFGE was run on a CHEF Mapper system (Bio-Rad). *b*, Analysis using a PFGE regime permitting resolution of comparatively short DNAs (up to about 0.5 Mbp). Bis-PNA (0.4 μM) and two pairs of methylation/restriction enzymes were used. Lanes 7, 8, results for the M·*Hha*I/*Hae*II pair (20 and 25 U, respectively); lanes 12, 13, results for the M·*Hpa*II/*Sma*I pair (10 and 15 U, respectively). Other lanes correspond to control experiments .

The M·*Hpa*II/*Sma*I enzymatic pair with the same bis-PNA yielded a different, clear-cut, rare cleavage pattern (*b* in the figure, lanes 12 and 13). Chromosomes I, III, VI and IX are not cut by the M·*Hpa*II/*Sma*I system, but strong new bands appear as a result of quantitative cutting of larger chromosomes (IV and XVI, data not shown). Sequence data for chromosomes III and II show that each carries a unique binding site for our PNA (bold), overlapping with the *Hae*II restriction site (underlined) and *Hha*I methylation site (italicized): *GGCGC*<u>**T**</u>**TCCCCT**. No cutting sites are found in the known sequences of those chromosomes not cut in our experiments.

We have shown that short bis-PNA clamps used in conjunction with appropriate pairs of methylation and frequently cutting restriction enzymes can yield large fragments of genomic DNA. Although only pyrimidine bases can be used, that allows for 2^8=256 different sequences. Together with the large number of pairs of methylation and frequently cutting restriction enzymes available, a potentially very large number of rare-cutting strategies thus become available.

Alexei G. Veselkov
Vadim V. Demidov
Maxim D. Frank-Kamenetskii*
Center for Advanced Biotechnology and Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA
Peter E. Nielsen
Center for Biomolecular Recognition, Panum Institute, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark

1. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. *Science* **254**, 1497–1500 (1991).
2. Demidov, V. V., Yavnilovich, M. V., Belotserkovskii, B. P., Frank-Kamenetskii, M. D. & Nielsen, P. E. *Proc. natn. Acad. Sci. U.S.A.* **92**, 2637–2641 (1995) .
3. Egholm, M. *et al. Nucleic Acids Res.* **23**, 217–222 (1995).
4. Koob, M. & Szybalski, W. *Science* **250**, 271–273 (1990).
5. Strobel, S. & Dervan, P. *Nature* **350**, 172–174 (1991)
6. Ferrin, L. & Camerini-Otero, D. *Science* **254**, 1494–1497 (1991).

*To whom correspondence should be addressed.