Nucleic Acids Research, 1993, Vol. 21, No. 9 **2103–2107**

Sequence selective double strand DNA cleavage by Peptide Nucleic Acid (PNA) targeting using nuclease S1

Vadim Demidov, Maxim D.Frank-Kamenetskii*, Michael Egholm[1], Ole Buchardt[1] and Peter E.Nielsen[2,]*
Institute of Molecular Genetics, Russian Academy of Sciences, Kurchatov Square, 123182 Moscow, Russia, [1]Department of Organic Chemistry, The H.C. Ørsted Institute, Universitetsparken 5, DK-2100 Copenhagen Ø and [2]Department of Biochemistry B, The Panum Institute, Blegdamsvej 3c, DK-2200 Copenhagen N, Denmark

Received February 10, 1993; Revised and Accepted April 5, 1993

ABSTRACT

A novel method for sequence specific double strand DNA cleavage using PNA (peptide nucleic acid) targeting is described. Nuclease S1 digestion of double stranded DNA gives rise to double strand cleavage at an occupied PNA strand displacement binding site, and under optimized conditions complete cleavage can be obtained. The efficiency of this cleavage is more than 10 fold enhanced when a tandem PNA site is targeted, and additionally enhanced if this site is in trans rather than in cis orientation. Thus in effect, the PNA targeting makes the single strand specific nuclease S1 behave like a pseudo restriction endonuclease.

INTRODUCTION

Most restriction enzymes recognize quartet or sextet DNA sequences and only very few require octets, or in some unique cases dodecets for recognition. Furthermore, enzymes have been identified and isolated for only a small subset of all possible sequences within these constraints. In connection with the studies of large genomic DNA molecules in general and within the human genome project in particular a need has arisen to recognize and specifically cleave DNA molecules at more rarely occurring sites, *i.e.*, sites consisting of 8–16 base pairs. Thus efforts have been made to create artificial 'restriction nucleases', or to modify the procedure of using conventional restriction enzymes for this purpose. They include oligonucleotides, capable of binding sequence specifically to DNA *via* triplex formation, tagged with chemical (photochemical) groups or enzymes capable of cleaving DNA (1–5), or numerous modifications of the 'Achilles heel' general strategy (6–9). We now describe a quite different approach, which stems from the ability of homopyrimidine peptide nucleic acid (PNA) (10–12) to locally unwind the DNA duplex by strand displacement rendering the displaced strand sensitive to S1 nuclease (10,13,14). We have now found conditions under which S1 nuclease induces double strand DNA breaks at the site of PNA binding with high yields. This PNA directed and provoked double strand DNA cleavage by S1 nuclease is especially efficient when two adjacent PNA sites are targeted yielding quantitative double-strand breaks.

MATERIALS AND METHODS

PNAs were synthesized as described (10,11). Plasmids containing the target sequences were obtained by cloning of the appropriate oligonucleotides into the vector pUC19 or pBluescriptKS$^+$. To obtain pT10, the 16-mers 5′-GATCCT$_{10}$G and 5′-GATCCA$_{10}$G were cloned into the *Bam*H1 site (pUC19). pT9C and pT9CT9C were obtained by cloning 5′-TCGACT$_4$CT$_5$G and 5′-TCGACA$_5$GA$_4$G into the *Sal*I site (pUC19), and pT9CA9GKS was obtained by cloning 5′-TCGACT$_4$CT$_5$G and 5′-TCGACA$_5$GA$_4$G into the *Sal*I site of pBluescriptKS$^+$. The sequences of the inserts are:

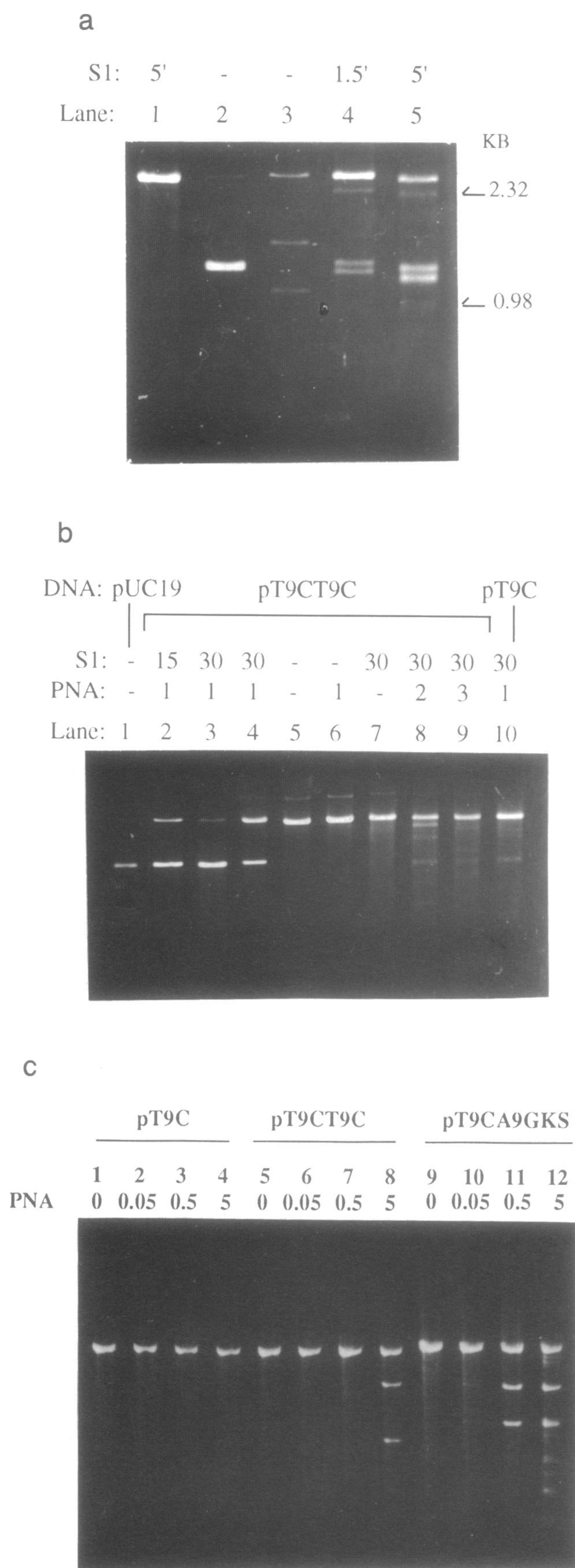
pT10: GGATCCTTTTTTTTTT(GGATCC)
pT9C: GGTCGACTTTTCTTTTT(GGATCC)
pT9CT9C: GGTCGACTTTTCTTTTTGTCGACTTTTCTTTTT(GGATCC)
pT9CA9GKS: GGTCGACAAAAAGAAAAGTCGACTTTTCTTTTT(GGATCC)

E. coli JM103 was used as host in all cases, and transformations and isolation of clones were done by standard techniques (15). Plasmids were purified by boyant density centrifugation in CsCl gradients. The plasmids were linearized with *Cfr*10I or *Sca*I restriction enzyme in the unique site.

Complexes between PNA and plasmid DNA were formed by incubating about 0.1 μg of the linearized plasmid with the desired amount of PNA in 10 μl of the TE buffer (10 mM Tris–HCl; 1 mM EDTA, pH 7.4) for 60 min at 37°C.

The S1 nuclease digestion was performed by adding 2 μl of 5×S1 buffer (250 mM NaAc; 1 M NaCl; 10 mM ZnSO$_4$; 0.5% of glycerol, pH 4.6) and S1 nuclease (Sigma or Boehringer Mannheim) to the above mixture and incubating for the desired time at room temperature or 37°C. The reaction was terminated by adding 1 μl of 0.5 M EDTA and cooling to −20°C. Electrophoresis was performed in 1% agarose gels in the TBE

* To whom correspondence should be addressed



buffer and DNA bands were subsequently visualised by staining with ethidium bromide.

3'-^{32}P-labelled DNA fragments were prepared using plasmids cut with restriction enzymes *Ecor*RI and *Pvu*II and labelling with α-^{32}P-ATP using the Klenow fragment of *E. coli* DNA polymerase (14,15).

RESULTS AND DISCUSSION

We first tried the direct approach. We used the *Cfr*101 linearized pT10 plasmid carrying dA_{10}/dT_{10} insert in the unique *Bam*H1 site in the polylinker, complexed it with PNA H-T_{10}-$LysNH_2$ and subjected the complex to treatment with S1 nuclease. Fig. 1a shows that after treatment with 145 units of S1 nuclease for 1.5 or 5 min at room temperature a significant fraction of DNA was cut into well defined fragments. The mobility by gel electroforesis of these fragments was very close to the mobility of fragments obtained by cuttting with the restriction enzyme *Bam*H1. These data clearly show that the S1 nuclease cleaves the DNA duplex at the site of PNA binding. Similar results were obtained for the PNA H-T_5CT_4-$LysNH_2$ and H-$T_2CT_2CT_4$-$LysNH_2$ and the corresponding pUC19 derived plasmids carrying the d(A_5G-A_4)/d(T_4CT_5) or d($A_2GA_2GA_4$)/d($T_4CT_2CT_2$) inserts. The cross experiments showed that the targeting was sequence specific: only complementary PNAs mediated cutting of the targets. Note, however, that two very week bands are seen in lanes 4 and 5, which correspond to the lengths 2.32 and 0.98 (Fig.1*a*). These bands are ascribed to weak binding of PNA T_{10}-$LysNH_2$ to the intrinsic pUC19 site d(T_7GT_3)/d(A_3CA_7), which is 0.37 kb apart from the *Cfr*10I restriction site. Fully consistent with previous results (14,16), these data show that even one mismatch dramatically decreases the affinity of PNA for the DNA target.

Figure 1. (**a**) Site-specific S_1 nuclease digestion of the pT10 plasmid linearized with *Cfr*10I restriction enzyme and complexed with PNA H-T_{10}-$LysNH_2$ (20 μM). Lane 1: control experiment; preincubation of the DNA with PNA was performed for 20 hours under conditions unfavorable for complex formation (200 mM of NaCl). Lane 2: the reference band obtained by digestion with *Bam*H1 restriction enzyme. Lane 3: reference bands obtained by the digestion with *Bgl*1 restriction enzyme. Lanes 4,5: the results of S1 nuclease digestion of the DNA-PNA complex 15 or 30 min, respectively. 10 U/μl of S1 was used in these experiments. (**b**) Site-specific S1 nuclease digestion of the pT9CT9C plasmid linearized by *Cfr*10I and complexed with PNA H-T_5CT_4-$LysNH_2$ (20 μM). The S1 digestion was for 30 min except for the sample of lane 2 which was digested for 15 min. Lane 1: pUC19 cleaved with *Cfr*10I and *Bam*H1. Lanes 2–4: site-specific digestion by the S1 nuclease of the linearized pT9CT9C plasmid complexed with PNA H-T_5CT_4-$LysNH_2$ (20 μM). In lane 4 preincubation was performed in the presence of Na^+ (50 mM). Lanes 5–7 are controls without PNA (lanes 5 & 7) and/or without S1 treatment (lanes 5 & 6). Lane 8: the same experiment as in lane 3, but using PNA H-T_{10}-$LysNH_2$ (20 μM) instead of H-T_5C-T_4-$LysNH_2$. Lane 9: the same experiment as in lane 3, but using PNA H-T_2C-T_2CT_4-$LysNH_2$ instead of H-T_5CT_4-$LysNH_2$. Lane 10: the same experiment as in lane 3, but using the pT9C plasmid instead of the pT9CT9C. In the header PNA H-T_5CT_4-$LysNH_2$ is labelled as 1, H-T_{10}-$LysNH_2$ as 2 and H-T_2CT_2C-T_4-$LysNH_2$ as 3. 2 U/μl of S1 was used in these experiments. The S1 units in each experiment are indicated in the header. (**c**) Site-specific S1 nuclease cleavage of the plasmids pT9C, pT9CT9C and pT9CA9GKS (linearized with *Sca*I) targeted with PNA T_5CT_4-$LysNH_2$. Lanes 1–4: pT9C; lanes 5–8: pT9CT9C; lanes 9–12: pT9CA9GKS. Lanes 1,5 & 9: no PNA; lanes 2,6 & 10: 0.5 μM; lanes 3,7,11: 5 μM; lanes 4,8,12: 50 μM. The samples were treated as above using 1 U/μl of S1 and 15 min incubation at 37°C.

We suggest that the enzyme first digests the displaced strand (as shown by probing experiments (10,16 and *vide infra*)), then to some extent enlarges the gap after which the opposite strand becomes a substrate for the enzyme. As a result a double strand break is created. The PNA-mediated digestion with the S1 nuclease leads to the clear-cut doublet However, whereas the two fragments are poorly resolved after digestion with *Bam*H1, . This most probably reflects a non-symmetrical widening of the gap by the S1 nucleasee as well as digestion from the ends of the *Cfr*10I site. With longer treatment the downward shift of the doublet becomes noticeable (see lane 5 in Fig. 1a). Thus, quantitative cleavage is clearly accompanied by a truncation of the fragments (by about one hundred base pairs). Although the above results were quite promising, they indicated that to obtain an artificial 'restriction enzyme' with the precision of cutting comparable to that of natural restriction enzymes the original approach had to be modified.

We therefore used the insert: 5'-$A_5GA_4GTCGACA_5GA_4$, cloned into the *Sal*1 site of the pUC19 plasmid as a target. This plasmid, which contains two binding sites for PNA-T_5C-T_4-$LysNH_2$ separated by six base pairs, was designated pT9C-T9C. We expected that the strand displacement at the two T_5CT_4 sites would lead to an opening of the entire region, including the sequence GTCGAC, thereby providing an easy substrate for the S1 nuclease in both strands (Fig. 2).

Fig. 1b shows that subjecting the pT9CT9C plasmid, linearized by *Cfr*10I and complexed with PNA H-T_5CT_4-$LysNH_2$, to 30 units of S1 nuclease for 5 min at room temperature results in conversion of the full-length DNA molecules into half-length fragments. The band thus obtained has the same position and width as the band, which is created by *Bam*H1 cleavage (cf. lanes 1 and 3 in Fig. 1*b*). Only careful inspection of the original photographs reveals very weak digestion of the pT9CT9C plasmid in the presence of PNA H-T_{10}-$LysNH_2$ or H-T_2CT_2C-T_4-$LysNH_2$ under conditions which result in quantitative cutting of the plasmid in the presence of PNA H-T_5CT_4-$LysNH_2$ (see lanes 8,9 in Fig. 1b). This emphasizes a remarkable sequence specificity of our artificial 'restriction enzyme'. Moreover, under the much milder S1-nuclease treatment necessary to generate the data of Fig. 1b with pT9CT9C, the yield of double-stranded breaks in the pT9C plasmid was extremely low (see Fig.1b, lane 10).

An even more favorable situation arises if the two PNA targets are on opposite strands (plasmid pT9CA9GKS). In this case PNA directed S1 double strand cleavage is 10 fold more efficient (Fig. 1c, lanes 11−12) as compared to the situation having both PNA targets on the same strand (Fig. 1c, lanes 7−8). The three possible senarios are sketched in Fig. 2, which also indicated the cutting sites by S1. More experiments are needed to determine the flexibility of the system in terms of the gap between the PNA targets and the size of these targets.

The extent of double stranded cleavage is dependent on both the concentration of PNA and on the digestion time as shown in Fig. 3, and it is particularly noteworthy that higher than optimal concentrations of PNA give rise to cleavage at multiple sites. This is ascribed to the binding of PNA to lower affinity, mismatch sites. The results of Fig. 3 also show that mung bean nuclease, another single strand specific nuclease, can substitute for nuclease S1.

Quantitation of the S1 digestion in a time course experiment (Fig. 4) showed that under conditions that resulted in total digestion of the full length plasmid a parallel formation of the fragments (at a constant ratio) corresponding to the site specifically cleaved product is taking place (Fig. 4b). Under these

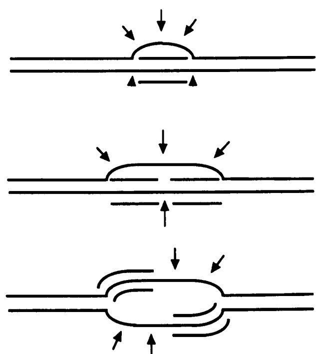


Figure 2. Schematic model of the three scenarios proposed for PNA targeted S1 double stranded cleavage of DNA. **a)** Single target. **b)** Double target in cis orientation. **c)** Double target in trans orientation. S1 attack is indicated by arrows. The binding of two PNA per DNA strand is inferred from the stoichiometry found upon binding of PNA to homopurine oligonucleotides (11, 12), and from from chemical probing experiments using $KMnO_4$ and dimethylsulfate. The latter experiments show that N7 of guanine in the target DNA is protected upon strand displacement binding of PNA to A/G mixed targets like the one found in pT9C (manuscript in preparation).

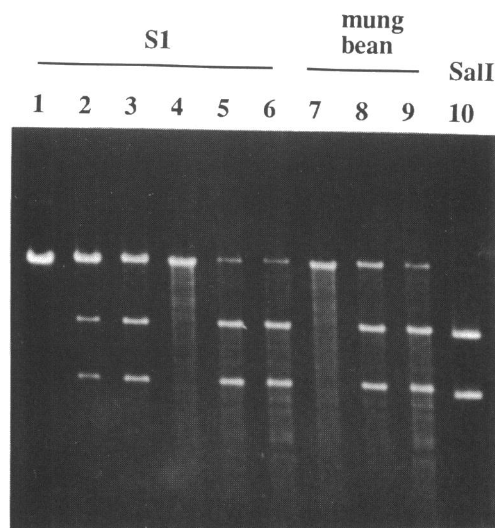


Figure 3. Dependence of double strand cleavage on PNA concentration and digestion time. All samples contained 50 μg/ml pT9CA9GKS plasmid DNA linearized with *Sca*I in 10 μl TE buffer and the following concentrations of PNA H-T_5CT_4-$LysNH_2$: lanes 1,4 & 7: None;lanes 2,5 & 8: 10 μM; lanes 3,6 & 9: 20 μM. After incubation at 37°C for 60 min, 2 μl 5× concentrated S1-buffer (1 M NaCl, 250 mM NaAc, pH 4.6, 10 mM $ZnSO_4$) was added, and the samples were treated as follows: Lanes 1−3: 10 U of S1, 10 min at 37°C; lanes 4−6: 10 U of S1, 30 min at 37°C; lanes 7−9: 30 U of mung bean nuclease, 10 min at 37°C. Lane 10 is a control containing pT9CA9GKS treated with ScaI and SalI (which cleaves at the PNA target site).

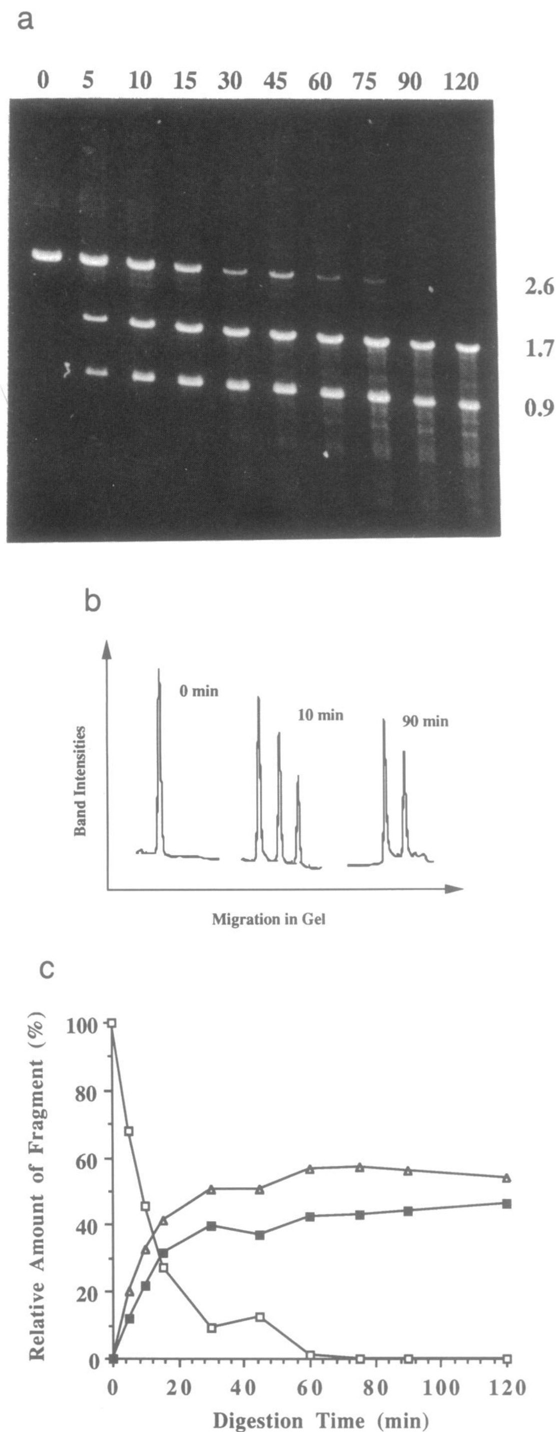


Figure 4. Time course of the PNA targeted S1 digestion of pT9CA9GKS. 0.5 μg pT9CA9GKS was incubated with 20 μM PNA H-T_5CT_4-$LysNH_2$ in 10 μl TE buffer for 60 min at 37°C. 2 μl 5×S1 buffer and 30 U nuclease S1 were added and the sample was incubated at 37°C for the times indicated. The samples were subsequently analysed by gel electrophoresis in agarose stained with ethidium bromide (**a**) and the DNA bands were quantified by densitometric scanning of the photographs (**b**), and plotted (**c**). □: linearised plasmid (2.6 kbp); △: 1.7 kbp fragment; ■: 0.9 kbp fragment. Relative amount is based on staining intensity.

conditions only minor cleavage at non target sites is detectable, and from the results shown in Fig. 4, we estimate the unspecific cleavage to be less than 10%.

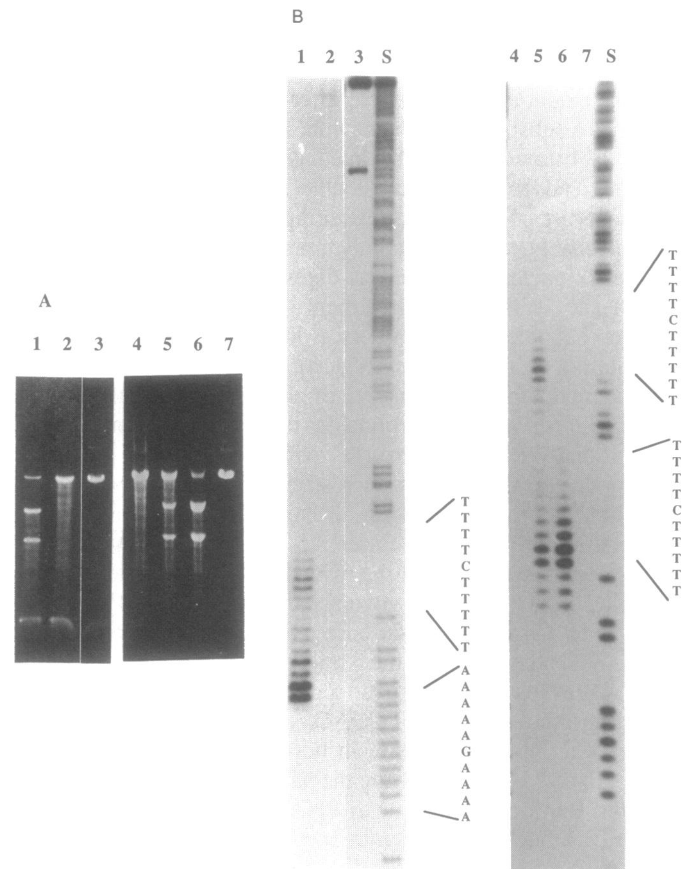


Figure 5. Sequence mapping of the double stranded DNA cleavage. A mixture of the pT9CA9GKS plasmid (0.5 μg linearized with *Sca*I) and a 3′-^{32}P-endlabeled *Pvu*II/*Eco*RI fragment of either pT9CA9GKS (lanes 1–3) or pT9CT9C (lanes 4–7)) was targeted with PNA H-T_5CT_4-$LysNH_2$ (20 μM) and treated with S1 (1 or 3 U/μl for 30 min at 37°C). Aliquots of each sample were analyzed both by **a)** electrophoresis in 1% agarose gels (ethidium stained), or by **b)** electrophoresis in polyacrylamide sequencing gels (visualized by autoradiography). Lane 1: 20 μM PNA. Lane 2: no PNA. Lane 3: no PNA, no S1. S: A+G sequence reaction of pT9CA9GKS. Lane 4: no PNA. Lane 5: 20 μM PNA, 1 U/μl S1. Lane 6: 20 μM PNA, 3 U/μl S1. Lane7: no S1, no PNA control. S: A+G sequence reaction of pT9CT9C.

In order to determine the extent of S1 digestion at the PNA site resulting upon double strand cleavage of the optimal DNA target, we performed an experiment in which the target was present both as an *Sca*1 linearized pT9CA9GKS plasmid as well as a ^{32}P-endlabelled *Eco*R1-*Pvu*II fragment from this plasmid. Following S1 digestion the analysis was performed both by electrophoresis in agarose gels and in polyacrylamide sequencing gels (Fig. 5). Under conditions that result in ~90% double strand cleavage, a distribution of cleavage sites within the predicted single stranded region of the double PNA target is observed (Fig. 5, lanes 1). However, the strongest bands are seen at the 3′-end of the loop showing that most of the loop has been digested.

It was not possible to do the analogous experiment looking at the opposite strand since this requires 5′-end labeling of the DNA fragment, and this label is lost during the experiment due to S1 digestion of the breathing end. However, by using the pT9C-T9C plasmid carrying the two inserts in cis and conditions which result in >90% double strand cleavage of the pT9CA9GKS plasmid we could show that the digestion is almost exclusively

confined to the DNA of the proposed single stranded loop, and the digestion extends only very weekly 2–3 bases into the proposed double stranded region (Fig. 5, lanes 5 & 6).

We conclude that this PNA directed double strand S1 cleavage will result in DNA where most of the single stranded loop has been digested, and thus the products will have overhangs corresponding to the PNA binding site. Depending on the orientation of the PNA sites 5′- or 3′- overhangs will be produced.

CONCLUSION

Our results demonstrate a novel strategy for sequence-selective cleavage of double stranded DNA. For cases where two closely positioned homo-purine/pyrimidine stretches (of 7–10 bases and preferably on opposite strands) can be identified this can be done just by synthesizing a pair of PNAs complementary to the purine strands of these two DNA regions. Provided that the strand displacement binding mode can be extended to include PNAs recognizing DNA sequences containing thymines and cytosines, this strategy could allow targeting and specific clevage of any desired sequence of 10–20 base pair. This is, however, not straightforward due to the fact that the high stability of PNA-DNA complexes consist of two PNA strands (presumably one 'Watson-Crick' and one 'Hoogsteen') and one DNA strand (Nielsen at al., in preparation). Thus in the case of mixed sequences *within* the PNA recognition sites the technology still needs development with respect to the sequence of the 'Hoogsteen-like' strand.

We foresee that PNA directed S1 targeting could be a valuable tool for DNA analysis in general, and for chromosome mapping, gene cloning and DNA based diagnostics in particular.

ACKNOWLEDGEMENTS

We thank Boris Belotserkovskii and Alexei Veselkov for help and stimulating discussions, and Karin Frederiksen for expert technical assistance. This work was supported by PNA Diagnostics, ISIS Pharmaceuticals and the Millipore Corporation.

REFERENCES

1. Francois, J.-C., Saison-Behmoaras, T., Barbier, C., Chassignol, M., Thuong, N.T., and Hélène, C. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9702–9706.
2. Perrouault, L., Asseline, U., Rivalle, C., Thuong, N.T., Bisagni, E., Giovannangeli, C., Le Doan, T. and Hélène, C. (1990) *Nature* **344**, 358–360.
3. Strobel, S.A. and Dervan, P.B. (1990) *Science* **249**, 73–75.
4. Pei, D., Corey, D.R. & Schultz, P.G. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9858–8962.
5. Beal, P.A. and Dervan, P.B. (1991) *Science* **251**, 1360–1363.
6. Hanvey, J.C., Shimuzu, M. and Wells, R.D. (1990) *Nucleic Acids Res.* **18**, 157–161.
7. Koob, H. and Szybalski, W. (1990) *Science* **250**, 271–273.
8. Strobel, S.A. and Dervan, P.B.(1991) *Nature* **350**, 172–174.
9. Ferrin, L.J. and Camerini-Otero, R.D. (1991) *Science* **254**, 1494–1497.
10. Nielsen, P.E., Egholm, M., Berg, R.H. and Buchardt, O. (1991) *Science* **254**, 1497–1500.
11. Egholm, M., Buchardt, O., Nielsen, P.E. and Berg, R.H. (1992) *J. Amer. Chem. Soc.* **114**, 1895–1897.
12. Egholm, M., Nielsen, P.E., Buchardt, O. and Berg, R.H. (1992) *J. Amer. Chem. Soc.* **114**, 9677–9678.
13. Cherny, D.Y., Belotserkovskii, B.P., Frank-Kamenetskii, M.D., Egholm, M., Buchardt, O., Berg, R.H. and Nielsen, P. E. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1667–1670.
14. Nielsen, P.E., Egholm. M., Berg, R.H. and Buchardt, O. (1993) *Anti Cancer Drug Design* (in press).
15. Maniatis, T., Fritsch, E. & Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Habor Laboratory, Cold Spring Harbor, NY (1982).
16. Nielsen, P.E., Egholm. M., Berg, R.H. and Buchardt, O. (1993) *Nucleic Acids Res.* **21**, 197–200.