# JMB

# An Experimental Study of Mechanism and Specificity of Peptide Nucleic Acid (PNA) Binding to Duplex DNA

## Heiko Kuhn[1], Vadim V. Demidov[1], Peter E. Nielsen[2] and Maxim D. Frank-Kamenetskii[1]*

[1]*Center for Advanced Biotechnology, Department of Biomedical Engineering, Boston University, 36 Cummington St Boston, MA 02215, USA*

[2]*Center for Biomolecular Recognition, Department for Biochemistry and Genetics Laboratory B, The Panum Institute, Copenhagen University, Blegdamsvej 3c DK-2200 Copenhagen N Denmark*

We investigated the mechanism and kinetic specificity of binding of peptide nucleic acid clamps (bis-PNAs) to double-stranded DNA (dsDNA). Kinetic specificity is defined as a ratio of initial rates of PNA binding to matched and mismatched targets on dsDNA. Bis-PNAs consist of two homopyrimidine PNA oligomers connected by a flexible linker. While complexing with dsDNA, they are known to form P-loops, which consist of a $[PNA]_2$-DNA triplex and the displaced DNA strand. We report here a very strong pH-dependence, within the neutral pH range, of binding rates and kinetic specificity for a bis-PNA consisting of only C and T bases. The specificity of binding reaches a very sharp and high maximum at pH 6.9. In contrast, if all the cytosine bases in one of the two PNA oligomers within the bis-PNA are replaced by pseudoisocytosine bases (J bases), which do not require protonation to form triplexes, a weak dependence on pH of the rates and specificity of the P-loop formation is observed.

A theoretical analysis of the data suggests that for (C + T)-containing bis-PNA the first, intermediate step of PNA binding to dsDNA occurs *via* Hoogsteen pairing between the duplex target and one oligomer of bis-PNA. After that, the strand invasion occurs *via* Watson-Crick pairing between the second bis-PNA oligomer and the homopurine strand of the target DNA, thus resulting in the ultimate formation of the P-loop. The data for the (C/J + T)-containing bis-PNA show that its high affinity to dsDNA at neutral pH does not seriously compromise the kinetic specificity of binding. These findings support the earlier expectation that (C/J + T)-containing PNA constructions may be advantageous for use *in vivo*.

© 1999 Academic Press

*Keywords:* peptide nucleic acid (PNA); homopyrimidine bis-PNA; PNA/ DNA complexes; Hoogsteen pairing; triplexes

*\*Corresponding author*

## Introduction

Peptide nucleic acid (PNA) is a synthetic DNA mimetic consisting of natural or modified nucleobases attached to a protein-like, polyamide backbone (see inset in Figure 1). Due to its biostability, high affinity to DNA and other properties, PNA has been finding numerous applications as a research, diagnostic and therapeutic tool (reviewed most recently by Uhlmann *et al.*, 1998).

Here, we focus on the strand-invasion complexes of homopyrimidine PNAs and double-stranded DNA (dsDNA) dubbed the P-loops. As Figure 1 shows, the two homopyrimidine PNA oligomers within the P-loop form a $[PNA]_2$-DNA triplex while the second DNA strand is displaced (Cherny *et al.*, 1993; Demidov *et al.*, 1993, 1994, 1995; Nielsen *et al.*, 1994; Wittung *et al.*, 1996, 1997; Uhlmann *et al.*, 1998). P-loops are quite different from the well-known $[DNA]_3$ or RNA-$[DNA]_2$ triplexes that are assembled when homopyrimidine DNA or RNA oligomers bind to corresponding sites on dsDNA (reviewed by Frank-Kamenetskii & Mirkin, 1995). The P-loop is characterized by extraordinary stability and high sequence-specificity of formation. In addition, its formation
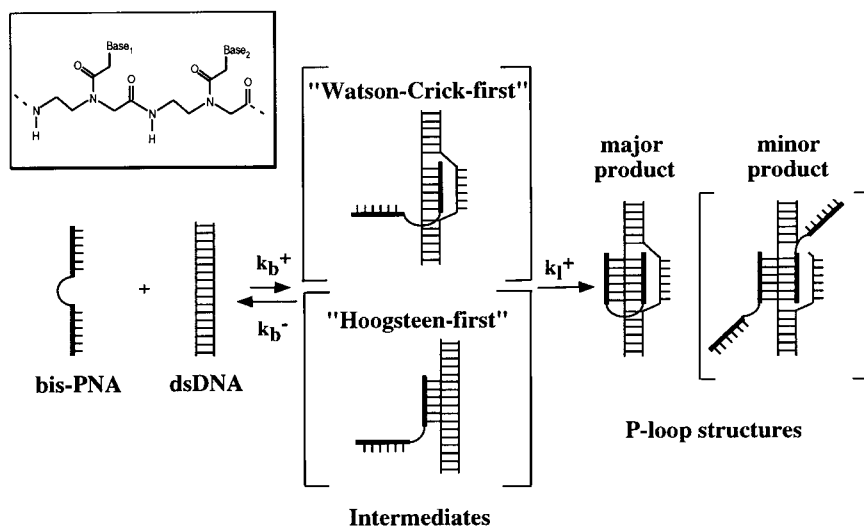
**Figure 1.** Two possible routes for the P-loop formation by bis-PNA inside dsDNA (Hoogsteen-first and Watson-Crick-first mechanisms). The inset shows the chemical structure of PNA.

locally opens the DNA double helix. These features make P-loops very attractive for applications (Demidov *et al.*, 1993; Møllegaard *et al.*, 1994; Veselkov *et al.*, 1996a,b; Izvolsky *et al.*, 1998; Bukanov *et al.*, 1998).

To enhance strand invasion efficiency, PNA clamps or bis-PNAs are often used (Figure 1). The bis-PNAs are two homopyrimidine PNA oligomers that are covalently connected *via* a flexible linker (Egholm *et al.*, 1995; Griffith *et al.*, 1995; Demidov *et al.*, 1995; Betts *et al.*, 1995). Linkage of two PNA oligomers offers the possibility to expediently design one PNA strand preferentially for Hoogsteen binding and the other PNA strand preferentially for Watson-Crick binding (Egholm *et al.*, 1995). This also reduces a trimolecular reaction of PNA to DNA binding to a bimolecular reaction, accelerating the PNA invasion. PNA strand invasion into dsDNA can be additionally accelerated by incorporation of positive charges into PNA or bis-PNA oligomers (Griffith *et al.*, 1995; Demidov *et al.*, 1996; Bentin & Nielsen, 1996; Veselkov *et al.*, 1996a; Kuhn *et al.*, 1998).

Experimental data and their analysis show that conditions exist under which high affinity of homopyrimidine PNAs to their target sites on dsDNA is accompanied by a high specificity of interaction (Demidov *et al.*, 1995, 1996, 1997; Kuhn *et al.*, 1998). It was concluded that the affinity of PNAs to their target sites is so high that PNA binding to correct and even to mismatched sites should be considered as virtually irreversible. Demidov *et al.* (1995, 1996) assumed that P-loop formation obeys a new principle of biomolecular recognition and takes place at two steps: a highly reversible first stage, which provides the specificity of binding, and a virtually irreversible second stage, which results in high binding affinity.

As Figure 1 shows, two simplified mechanisms of the strand invasion reaction are possible (Nielsen *et al.*, 1993; Demidov *et al.*, 1995). According to one mechanism, the first stage consists of a fluctuating opening of the DNA double helix and a transient formation of a PNA-DNA Watson-Crick duplex (the Watson-Crick-first mechanism). According to the other mechanism, the first stage consists of formation of an unstable PNA-$[DNA]_2$ triplex *via* Hoogsteen pairing (the Hoogsteen-first mechanism). Since the final result is the same for both mechanisms and since the unstable intermediates are very difficult to detect, the particular route of the P-loop formation remained elusive.

Here, we investigate the dependence of kinetic parameters of PNA binding to dsDNA on pH. A major parameter we determine is the kinetic specificity, $r$. According to a recent theoretical study by Lomakin & Frank-Kamenetskii (1998), pH-dependence of the $r$ value must be radically different for the two routes of PNA invasion into the double helix presented in Figure 1. The $r$ value is defined as a ratio of the initial rates of binding to matched and mismatched targets (Demidov *et al.*, 1995, 1997; Kuhn *et al.*, 1998; Lomakin & Frank-Kamenetskii, 1998). A theoretical analysis by Lomakin & Frank-Kamenetskii (1998) yielded a bell-shaped curve for the PNA kinetic specificity as a function of the apparent equilibrium stability constant $s$ for the first step of the two-step binding process. The particular meaning of the $s$ value depends on the mechanism assumed. For the Watson-Crick-first mechanism, the $s$ value is a relative stability constant determined by the ratio of the stability constant of the PNA/DNA duplex and the stability constant of the DNA/DNA duplex. For the Hoogsteen-first mechanism, the $s$ value is the stability constant of the intermediate PNA/dsDNA triplex formation.

Complexing of C-containing homopyrimidine oligonucleotides to a duplex DNA target is known to be strongly pH-dependent (reviewed by Frank-Kamenetskii, 1992). To form Hoogsteen bonds within triads, the third-strand cytosine bases must be protonated (see Figure 2). The $s$ value for elongation of the protonated triplex is expressed
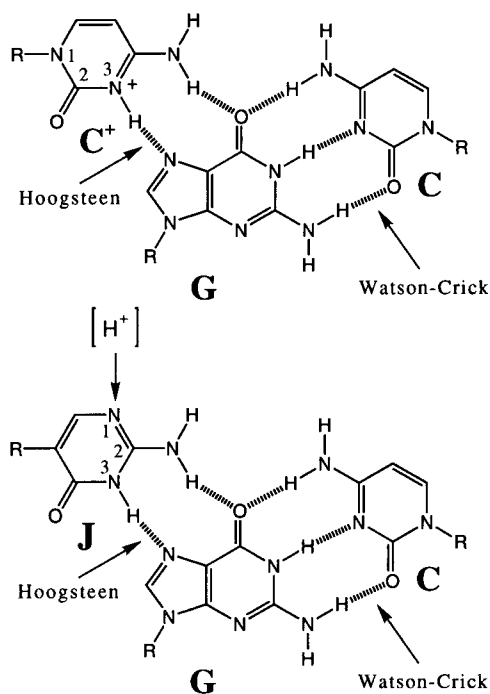
**Figure 2.** Chemical structures of $C^+ \cdot G$-C and $J(J^+) \cdot G$-C triads with Hoogsteen and Watson-Crick hydrogen bonds. Note that the J base can be protonated within the corresponding triad at the N1 position.

pH. Here, we present experimental data for binding of $(C + T)$-containing bis-PNA to dsDNA targets, which provide a strong evidence that P-loop formation occurs *via* the Hoogsteen-first route.

## Results

We have measured the initial rates of binding of $(C + T)$-containing bis-PNA I and $(C/J + T)$-containing bis-PNA II to complementary and single-mismatched targets at various pH. PNAs used and their target sites are shown in Figure 3. PNA II is identical with PNA I with the exception that it contains pseudoisocytosine (J base) instead of a cytosine in its N-terminal half. The J base carries a hydrogen atom at the N3 position allowing its Hoogsteen pairing with a guanine base without protonation (Figure 2). A bis-PNA containing J bases in one of the two PNA halves binds to dsDNA at pH 7.0 better than its $(C + T)$-containing analog (Egholm *et al.*, 1995).

In the binding reactions of PNAs I and II to target sites, the PNA concentration was kept at a large excess over the DNA concentration, and pseudo-first-order kinetics of binding were observed in all cases. While calculating the rate constants, we did not distinguish between the different PNA-dsDNA complexes observed for

*via* pH as:

$$s = (1 + 10^{pK_T - pH})^{1/r_H} \qquad (1)$$

where $pK_T$ is the p$K$ value for triplex protonation at the N3 position of the cytosine residue, and $r_H$ is the number of Hoogsteen base-pairs formed per one cytosine residue in the triplex (Lyamichev *et al.*, 1985, 1988; Frank-Kamenetskii, 1992). Equation (1) leads to a very steep dependence of the $s$ value on pH. Hence, for Hoogsteen binding of C-containing oligomers, the $s$ value can be easily changed in a monotonous way by more than an order of magnitude by varying the pH of the solution near the $pK_T$ value, which lies in the neutral range of pH (Frank-Kamenetskii, 1992).
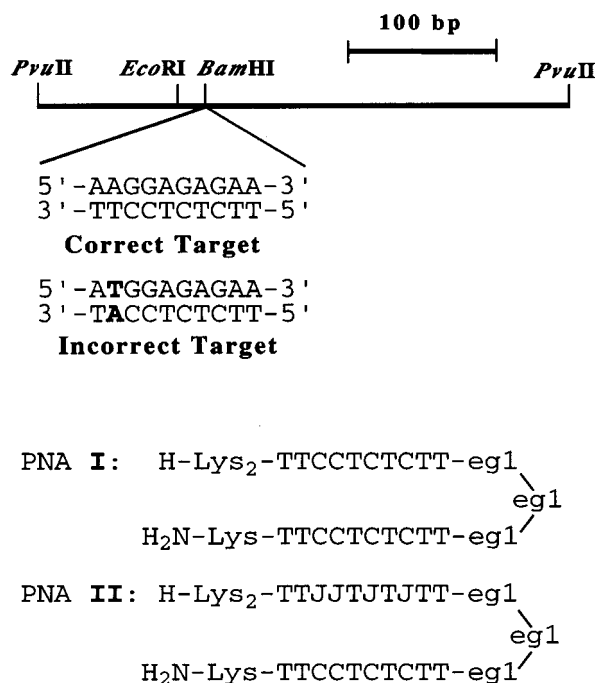
On the other hand, a theoretical analysis by Lomakin & Frank-Kamenetskii (1998) has shown that the $r$ value must exhibit a bell-shaped dependence on the stability constant $s$ for the intermediate complex (see Figure 5(b) and (c)). Therefore, we assume that in the case of the Hoogsteen-first mechanism we must observe a characteristic, bell-shaped dependence of kinetic specificity of the $(C + T)$-containing bis-PNA binding to dsDNA on pH within the neutral range of pH values. We also expect a significant change in the rate of binding of C-containing homopyrimidine PNA to dsDNA with pH. In contrast, there is no reason to expect any significant pH-dependence of kinetic specificity and PNA binding rate in the case of the Watson-Crick-first mechanism at neutral range of



**100 bp**

```
PvuII    EcoRI BamHI                          PvuII

5'-AAGGAGAGAA-3'
3'-TTCCTCTCTT-5'
  Correct Target

5'-ATGGAGAGAA-3'
3'-TACCTCTCTT-5'
  Incorrect Target
```

PNA **I:**  H-Lys$_2$-TTCCTCTCTT-eg1
                              eg1
        H$_2$N-Lys-TTCCTCTCTT-eg1

PNA **II:** H-Lys$_2$-TTJJTJTJTT-eg1
                              eg1
        H$_2$N-Lys-TTCCTCTCTT-eg1

**Figure 3.** DNA targets and bis-PNAs. The DNA constructs were *PvuII/PvuII* (338 bp) or *PvuII/EcoRI* (246 bp) fragments, isolated from pUC19 derivatives. PNAs are written from the N terminus to the C terminus using normal peptide conventions: H is a free amino group; NH$_2$ is a terminal carboxamide; Lys is the lysine residue; J denotes pseudoisocytosine (see Figure 2) and eg1 denotes the linker unit, 8-amino-3,6-dioxaoctanoic acid (Egholm *et al.*, 1995).
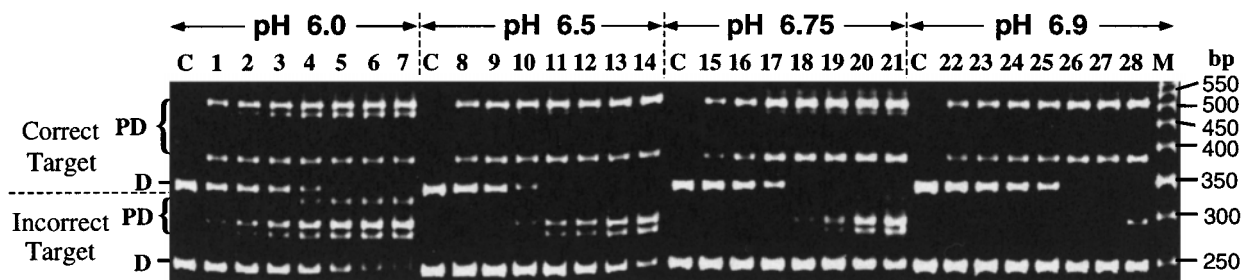
**Figure 4.** Gel-mobility-shift analysis of PNA I binding to the correct dsDNA target (a 338 bp long *Pvu*II/*Pvu*II DNA fragment) and to a single-mismatched (incorrect) target (a 246 bp long *Pvu*II/*Eco*RI DNA fragment). The binding was performed during the following periods of time: pH 6.0, 0.5, 1, 2, 4, 8, 15 and 30 minutes (lanes 1-7); pH 6.5, 1, 2, 4, 8, 15, 30, 60 minutes (lanes 8-14); pH 6.75, 2, 4, 10 and 30 minutes (lanes 15-18) and 1, 4, 8 hours (lanes 19-21); pH 6.9, 0.25, 0.5, 0.75, 1, 4, 8, 24 hours (lanes 22-28). C, controls without PNA; M, a 50 bp DNA ladder; D, is for free DNA; PD, PNA/DNA complexes.

correct and mismatched PNA binding (*vide infra*). Pseudo-first-order rate constants corresponding to the PNA initial binding rates for the non-equilibrium case of PNA binding were determined as described (Demidov *et al.*, 1995; Kuhn *et al.*, 1998). We used restriction fragments of derivatives of the pUC19 plasmid carrying the corresponding target sites (Figure 3).

Figure 4 shows the results of binding experiments performed with PNA I in the presence of matched and mismatched dsDNA targets in the same sample. Note that after binding of PNA I and PNA II to both correct and incorrect dsDNA targets we observed several (up to three) distinct bands corresponding to various PNA-dsDNA complexes. We have also observed similar complex patterns while targeting dsDNA with other bis-PNAs (Kurakin *et al.*, 1998; data not shown).

Two of these retarded bands may arise due to the different mobility of the PNA-dsDNA complex, depending on what side the bis-PNA linker passes the DNA strands at the target site. Another possibility consists in the formation of a strand-invasion complex by two different bis-PNA molecules (Figure 1). In support of the latter possibility, we observed that the relative intensity of one of the bands increases with increasing PNA concentration (data not shown). In a separate study (G. Hansen & P.E.N., unpublished results), the complexes mentioned above are assigned to different bands on the gels.

The data in Figure 4 demonstrate a transition from very poor (pH 6.0) to very good (pH 6.9) selectivity of targeting of dsDNA by PNA I. Indeed, complete PNA binding to the correct DNA target can be achieved at neutral pH virtually without any mismatched binding. In contrast, at pH 6.0 PNA I exhibits the same binding pattern in the case of correct and mismatched targets. Remarkably, the transition from very poor to very high specificity occurs within one unit of pH. Note that the quantitative values of kinetic specificity *r* in these experiments agree with the data in Figure 5, which were obtained in separate measurements of

kinetic constants of PNA I binding to correct and mismatched targets.

Figure 5(a) shows that the binding rate constant of PNA I to its complementary DNA target site changes by more than three orders of magnitude between pH 6.0 and 7.4. The most dramatic drop is observed above pH 6.8. At pH 6.8, 50 % of binding to the target site was reached after two minutes of incubation at 37 °C ($k = 0.35$ min$^{-1}$), while it took approximately two days to reach 50 % of binding at pH 7.4 ($k = 2.3 \times 10^{-4}$ min$^{-1}$) for the same PNA concentration. The binding rate to the mismatched target exhibits a similar drop. However, in this case the binding rate decreases already at pH values higher than 6.5. As a result, the value of kinetic specificity *r* of PNA I rises dramatically above pH 6.5, exhibiting a sharp maximum near pH 6.9. It reaches the value of approximately 200, then decreases again at higher pH values (Figure 5(b)). Above pH 7.2, the kinetic specificity could not be determined since the rate for PNA I binding to the mismatched DNA target became too low to be measured. Figure 5(c) shows the resulting plot of the kinetic specificity *r* of PNA I as a function of the stability constant *s*.

As expected, the binding rate of PNA II was significantly less pH-dependent than in case of PNA I. Figure 6(a) shows that binding of PNA II to the correct DNA target is only five times slower at pH 7.4 than at pH 6.0. In contrast, the binding rate of PNA I decreases about 10$^4$-fold over the same pH range. The binding rate of PNA II proved to be notably pH-dependent and more pronounced in the case of mismatched binding, which was not expected.

Figure 2 shows that the J base can be protonated at the N1 position, acquiring a net positive charge. As in case of cytosine bases, the corresponding p$K$ value for J base protonation may be significantly shifted upwards within a triplex as compared to the free base, and lie in the neutral range of pH. At slightly acidic pH, this could result in enhanced binding of protonated (C/J + T)-containing bis-PNA to negatively charged dsDNA due to the electrostatic attraction. The stabilizing effect of J base

(a)



(b)



(c)
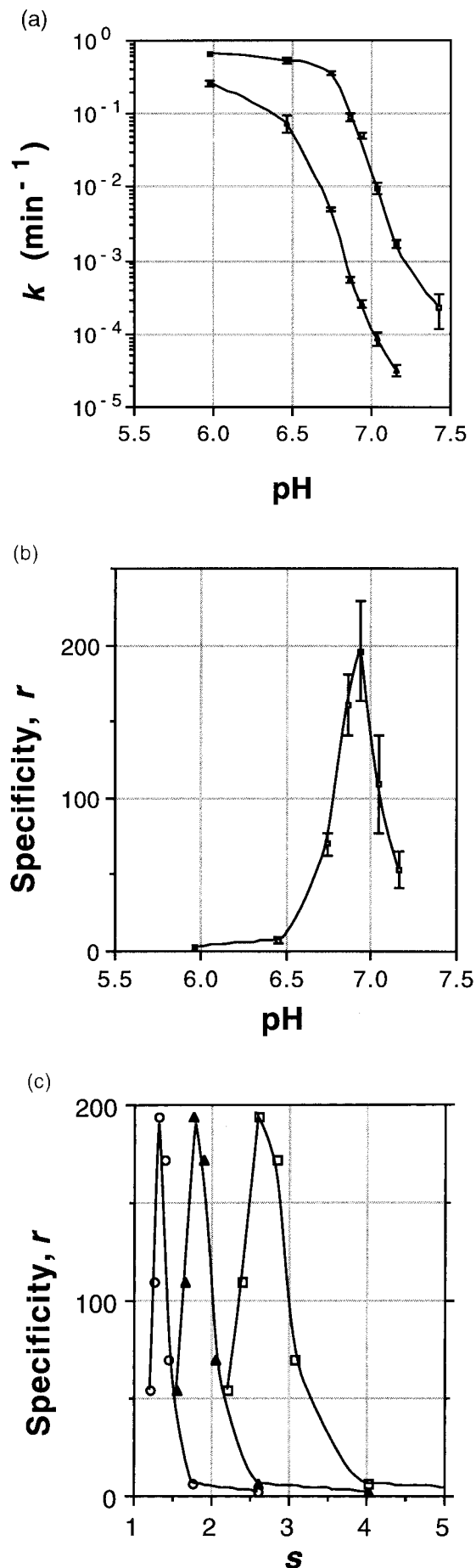


protonation should be more significant in the case of mismatched (and therefore weak) binding. However, we cannot exclude quite different and still unknown reasons for this effect, since we observed a similar pH dependence for a purely thymine-containing bis-PNA (data not shown). It should be emphasized again that in any case this pH dependence is much weaker than in case of PNA I.
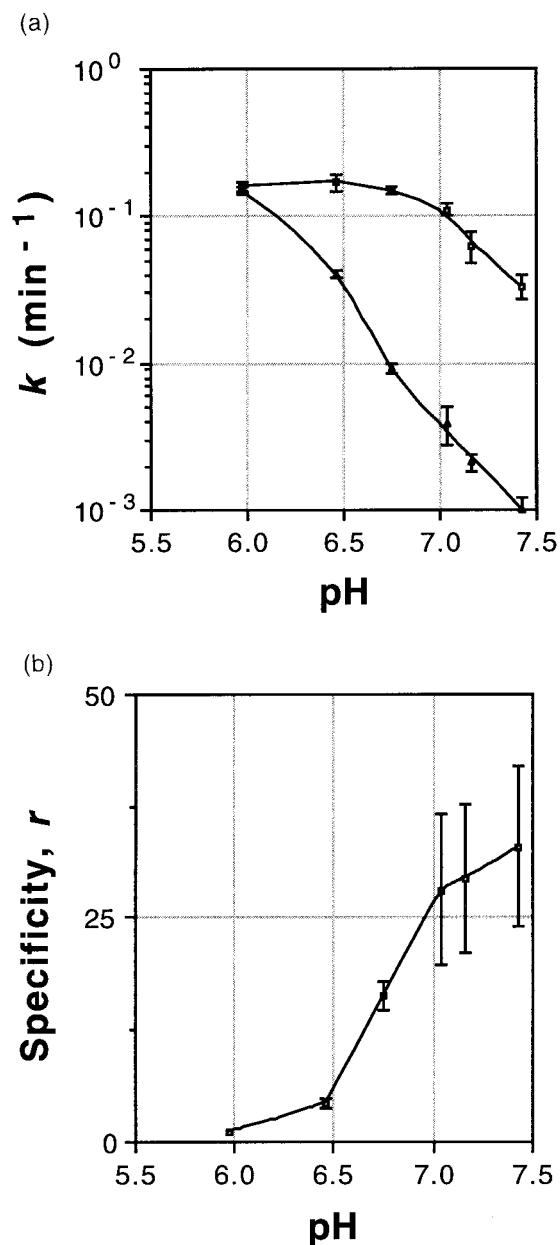
As for PNA I, specificity of PNA II binding decreases below pH 7.0. This may be due to the enhanced binding of this PNA because of possible protonation of J base, as discussed above. However, between pH 6.8 and 7.4, the pH-dependence of the PNA II binding rate to the mismatched DNA target is parallel with the binding rate to the matched DNA target. As a result, the kinetic specificity of PNA II is virtually constant in this pH range, being about 30 (Figure 6(b)).

## Discussion

Our experimental data demonstrate that the binding rate of a (C + T)-containing bis-PNA (PNA I) to its target site on dsDNA drops dramatically within the pH range of 6.5-7.4. Below pH 6.9, the binding rate of PNA I is close to the binding rate of PNA II, in which all cytosine bases in one of two bis-PNA oligomers are replaced by J bases. However, in the physiological range of pH (between 7.1 and 7.4), the binding rate for PNA I drops by a factor of $10^4$ while the rate for PNA II remains virtually constant for PNA binding with the correct DNA target. Therefore, bis-PNAs with mixed (C + T)-sequences are not good candidates for dsDNA targeting at neutral and slightly alkaline pH, and incorporation of J bases into bis-PNA is justified for their possible *in vivo* use. On the other hand, PNAs similar to PNA I may prove to be useful tools if the pH value can be appropriately chosen and accurately controled. Indeed, our data show that in a narrow range of pH between 6.8 and 7.0 a (C + T)-containing bis-PNA exhibits a high binding rate (Figure 5(a)) and an exceptionally high specificity of binding (Figure 5(b)).

The maximal kinetic specificity for (C + T)-containing PNA I is several times higher than the specificity value of its (C/J + T)-containing analog, PNA II (compare the data in Figures 5(b) and 6(b)). Further experiments are needed to find out

**Figure 5.** (a) The pseudo-first-order rate constant $k$ determined from PNA I/dsDNA binding experiments as a function of pH. The upper curve (squares) corresponds to the correct DNA target, the lower curve (triangles) corresponds to the single mismatched (incorrect) DNA target. The bars indicate standard deviations. (b) Kinetic specificity $r$ for PNA I as a function of pH. (c) Kinetic specificity $r$ for PNA I as a function of the stability constant $s$. The $s$ value is calculated using equation (1) with $pK_T$ 7.0 (left curve), 7.5 (middle curve) and 8.0 (right curve); $r_H = 2.5$.

(a)



(b)



**Figure 6.** (a) The pseudo-first-order rate constant $k$ determined from PNA II/dsDNA binding experiments as a function of pH. The assignment of curves to the matched and mismatched DNA targets is as described in the legend to Figure 5(a). (b) Kinetic specificity $r$ for PNA II as a function of pH.

Note that in this study a mismatched base was located near the end of the complex. The specificity is much higher (several hundreds and more) for mismatches located farther from the ends of the bis-PNA binding site (see Kuhn *et al.*, 1998). Indeed, in our experiments with PNA I and PNA II we failed to observe any complex formation when the mismatch was located at the fourth position from the end of the target site. For example, no mismatched binding was detected after incubation of 1.5 μM of PNA II at pH 7.0 for five days at 37 °C, while binding of this PNA under the same conditions with the correct DNA target was completed in about 30 minutes (data not shown). Therefore, the kinetic specificity value in this case is definitely higher than $10^3$. The lack of PNA binding to the DNA targets carrying mismatches in the middle of the target sites prevented us from quantitative analysis of such cases.

Our data shed a new light on the specific mechanism of the P-loop formation. We observed a very sharp, bell-shaped pH-dependence of the specificity of PNA I binding to dsDNA targets. In the two-step process of the PNA-DNA recognition, the kinetic specificity is determined mainly by the highly reversible first stage. Hence, no pH dependence is expected in case of the Watson-Crick-first mechanism because Watson-Crick pairing does not involve any protonation of nucleobases. In contrast, a strong pH-dependence is expected in case of the Hoogsteen-first mechanism (see Introduction).

The Hoogsteen-first mechanism can easily explain the strong dependence of the binding rate for PNA I on pH. In fact, a two-step process of binding of PNA to DNA (see Figure 1) can be described as a whole by a pseudo-first-order rate constant:

$$k \cong (k_b^+/k_b^-)k_1^+ P \tag{2}$$

where $P$ is the PNA concentration, $k_b^+$ and $k_b^-$ are the rate constants for the forward and reverse PNA binding to the DNA duplex, and $k_1^+$ is the rate constant for formation of $[PNA]_2$-DNA triplex (Demidov *et al.*, 1995, 1996; Lomakin & Frank-Kamenetskii, 1998). The equilibrium constant for triplex formation by a homopyrimidine oligonucleotide with dsDNA, which can be related to the $k_b^+/k_b^-$ ratio in equation (2) for the Hoogsteen-first mechanism of PNA binding, strongly depends on pH (Maher *et al.*, 1990). Therefore, the Hoogsteen-first mechanism is expected to lead to a steep pH-dependence of the binding rate $k$ for a (C + T)-containing PNA. In contrast, for the Watson-Crick-first mechanism, the $k_b^+/k_b^-$ ratio is obviously pH-independent in the neutral pH range. The $k_1^+$ value corresponds in this case to the association rate constant for triplex formation of PNA with the PNA-DNA heteroduplex. No direct data are available to that matter. Turning again to the much better studied case of triplex-forming oligonucleotide, the rate constant of its association with dsDNA is

whether the lower maximal kinetic specificity for PNA II is a rule when C bases are replaced by J bases in one of two PNA oligomers of bis-PNA. Note that the $r$ value of 30 already provides good discrimination for dsDNA targeting: almost complete binding to the correct target can be reached for a certain incubation time without noticeable binding to the mismatched target (Demidov *et al.*, 1997). Nevertheless, higher specificity is desirable, especially for PNA targeting of long genomic DNA.

known to be a weak function of pH. It decreases by two to three times between pH 6.8 and pH 7.2 (Maher *et al.*, 1990). In contrast to this data, we have observed in this range a dramatic change in the PNA binding rate by two orders of magnitude. Thus, our data on the pH-dependence of the $k$ value can by no means be explained on the basis of the Watson-Crick-first mechanism.

One more line of evidence in favor of the Hoogsteen-first mechanism stems from the fact that before leveling off, the binding rate of PNA I changes about $10^4$ times, while pH changes by one unit (see Figure 5(a)). It is not accidental that the value of exponent is close to the number of cytosine bases (4) in the Hoogsteen half of bis-PNA I. This should be expected on the basis of equation (2) for the Hoogsteen-first mechanism because the pH-dependent factor in this expression for $pH < pK_T$ must be:

$$k_b^- \cong k_o \, 10^{m \, pH} \qquad (3)$$

where $k_o$ is a constant and $m$ is the number of nucleobases participating in the formation of protonated triads (Lyamichev *et al.*, 1985, 1988; Frank-Kamenetskii, 1992). For a $(C + T)$-containing 10-meric mono-PNA with two cytosine bases, the binding rate changes 100-fold per one unit of pH in the neutral pH range, also in agreement with equation (3) (data not shown).

Using equation (1), we can plot the kinetic specificity values as a function of the stability constant $s$ corresponding to the PNA-dsDNA transient Hoogsteen pairing. To do so we need the $pK_T$ value for the cytosine protonation within the PNA-dsDNA triplex. The crystallographic data by Betts *et al.* (1995) for the [PNA]$_2$-DNA triplex show that cytosine bases involved in the Hoogsteen pairing are protonated at pH 8.5. The kinetic "titration curve" for the matched PNA-DNA complex in Figure 5(a) indicates that the $pK_T$ value for cytosine bases within the triplex is close to or even larger than 7. This $pK_T$ value correlates well with the value reported for (DNA)$_3$ triplexes (Howard *et al.*, 1964; Lyamichev *et al.*, 1985; Frank-Kamenetskii, 1992; Wilson *et al.*, 1994; Leitner *et al.*, 1998). To calculate the stability constant $s$ using equation (2), we chose the $pK_T$ values in the range between 7.0 and 8.0. Figure 5(c) shows the resulting plots of the kinetic specificity $r$ as a function of $s$.

Although our data cannot be quantitatively compared with the theoretical predictions by Lomakin & Frank-Kamenetskii (1998), the parallel between the theoretical predictions on the one hand and our experimental data on the other is striking. Indeed, the curves in Figure 5(c) are very similar to the bell-shaped curves simulated by Lomakin & Frank-Kamenetskii (1998). Thus, our data strongly indicate that in the case of bis-PNAs we studied here the process proceeds *via* the Hoogsteen-first route. Furthermore, it is important to note that according to the theoretical analysis, in case of a two-step binding process both high affinity and

high specificity can be reached in a concerted way (Demidov *et al.*, 1997; Lomakin & Frank-Kamenetskii, 1998).

Equation (2) describes the case of maximum achievable kinetic specificity and is valid only under quasistationary conditions of PNA binding, when the following inequalities are met (Demidov *et al.*, 1995, 1996; Lomakin & Frank-Kamenetskii, 1998):

$$k_b^- \gg k_b^+ P; \quad k_1^+ \qquad (4)$$

The proposed sequence of elementary events of PNA invasion with the Hoogsteen-first and the Watson-Crick-second steps explains the nature of a large ratio of kinetic constants for the first and the second steps of the process, which, according to Lomakin & Frank-Kamenetskii (1998), is absolutely necessary for high kinetic specificity of the overall process. The PNA-dsDNA triplex formation is not associated with any substantial kinetic barrier. In contrast, the second step in the process of the P-loop formation must be associated with a significant kinetic barrier. Indeed, to convert the PNA-dsDNA triplex into [PNA]$_2$-DNA triplex, the purine DNA strand in the target site must exchange its Watson-Crick DNA partner for the Watson-Crick PNA partner. The height of the barrier must increase with the ionic strength.

We conclude that, although other modes of homopyrimidine PNAs binding to dsDNA are possible, our data strongly indicate that at neutral pH C-containing PNAs form the P-loops *via* the Hoogsteen-first, Watson-Crick-second two-step process.

## Materials and Methods

### PNAs

Bis-PNAs used in this study (see Figure 3) were synthesized by standard methods, purified by HPLC, and characterized by MALDI-TOF mass spectrometry as described (Egholm *et al.*, 1995; Christensen *et al.*, 1995). PNA concentrations were determined spectrophotometrically at 260 nm using the following molar extinction coefficients: 8800 (T), 7300 (C), and 3000 (J) $M^{-1}cm^{-1}$.

### Plasmids

Plasmids containing the target sequences were constructed by cloning the appropriate oligonucleotides into the polylinker of the pUC19 vector (see Figure 3 for details). They were purified by QIAGEN Midi Plasmid Kit and sequenced. For the gel-mobility-shift assay, DNA fragments containing the target site were prepared by cutting the plasmids with the *Pvu*II restriction enzyme or with the restriction enzymes *Pvu*II and *Eco*RI. Digested plasmids were stored at ~1 µg/µl in TE buffer (pH 7.4)

### PNA-DNA binding

Experiments of PNA binding to dsDNA fragments were performed in pre-siliconized tubes at 37 °C in 20 mM Na-phosphate buffer solutions at various pH

between 6.0 and 7.4. The PNA concentration in the reaction mixture was 1.5 μM, and the DNA concentration was ~0.01 μM so that the PNA was in large excess over DNA. The pH values of the reaction mixtures were measured with a Micro-pH Combination Electrode (Orion, Model 9810). At selected time points the reaction aliquots were quenched by the addition of 1 M NaCl so that the total $Na^+$ concentration was 150 mM; the aliquots were immediately stored at $-20\,^\circ C$.

The gel-mobility-shift experiments were performed in non-denaturing 10 % polyacrylamide gels in $1 \times$ TBE buffer (pH 8.0). The electrophoresis was run at 250 V for four hours at room temperature. The gels were then stained with ethidium bromide with subsequent quantification of the intensities of all bands. The intensity of the faster migrating band gives the quantity of DNA free of PNA, whereas slower migrating bands correspond to complexes which retard due to P-loop formation (Peffer *et al.*, 1993). By measuring the intensities of these bands, normalized with their integral intensity, one can follow the kinetics of the PNA binding to DNA. Quantitative analysis was performed after scanning with a CCD camera using the IS-1000 digital imaging system and image analysis software (Alpha Innotech Corporation). The kinetic binding experiments at each indicated pH value were performed in triplicate. Mean values and standard deviations were determined by routine statistical procedures. The curve-fitting procedure involved computer graphical data analysis using Cricket Graph software.

### Kinetic specificity

Kinetic specificity $r$ was defined as a ratio of the initial rates $k$ of PNA binding to matched and mismatched targets (Demidov *et al.*, 1995, 1997; Kuhn *et al.*, 1998; Lomakin & Frank-Kamenetskii, 1998):

$$r = k_{match}/k_{mismatch}$$

The initial binding rates were determined according to equations (1) and (2) by Kuhn *et al.* (1998).

## References

Bentin, T. & Nielsen, P. E. (1996). Enhanced peptide nucleic acid binding to supercoiled DNA: possible implications for DNA ''breathing'' dynamics. *Biochemistry*, **35**, 8863-8869.

Betts, L., Josey, J. A., Veal, J. M. & Jordan, S. R. (1995). A nucleic acid triple helix formed by a peptide nucleic acid-DNA complex. *Science*, **270**, 1838-1841.

Bukanov, N. O., Demidov, V. V., Nielsen, P. E. & Frank-Kamenetskii, M. D. (1998). PD-Loop: a complex of duplex DNA with an oligonucleotide. *Proc. Natl Acad. Sci. USA*, **95**, 5516-5520.

Cherny, D. Y., Belotserkovskii, B. P., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., Berg, R. H. & Nielsen, P. E. (1993). DNA unwinding upon strand-displacement binding of a thymine-substituted polyamide to double-stranded DNA. *Proc. Natl Acad. Sci. USA*, **90**, 1667-1670.

Christensen, L., Fitzpatrick, R., Gildea, B., Petersen, K. H., Hansen, H. F., Koch, T., Egholm, M., Buchardt, O., Nielsen, P. E., Coull, J. & Berg, R. H. (1995). Solid-phase synthesis of peptide nucleic acids (PNAs). *J. Peptide Sci.* **3**, 175-183.

Demidov, V., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O. & Nielsen, P. E. (1993). Sequence selective double strand DNA cleavage by PNA targeting using nuclease S1. *Nucl. Acids Res.* **21**, 2103-2107.

Demidov, V. V., Cherny, D. I., Kurakin, A. V., Yavnilovich, M. V., Malkov, V. A., Frank-Kamenetskii, M. D., Sönnichsen, S. H. & Nielsen, P. E. (1994). Electron microscopy mapping of oligopurine tracts in duplex DNA by peptide nucleic acid targeting. *Nucl. Acids Res.* **22**, 5218-5222.

Demidov, V. V., Yavnilovich, M. V., Belotserkovskii, B. P., Frank-Kamenetskii, M. D. & Nielsen, P. E. (1995). Kinetics and mechanism of polyamide (''peptide'') nucleic acid binding to duplex DNA. *Proc. Natl Acad. Sci. USA*, **92**, 2637-2641.

Demidov, V. V., Frank-Kamenetskii, M. D. & Nielsen, P. E. (1996). Complexes of duplex DNA with homopyrimidine peptide nucleic acid (PNA): new principle of biomolecular recognition. In *Biomolecular Structure and Dynamics* (Sarma, R. H. & Sarma, M. H., eds), vol. 2, pp. 129-134, Adenine Press, NY.

Demidov, V. V., Yavnilovich, M. V. & Frank-Kamenetskii, M. D. (1997). Kinetic analysis of specificity of duplex DNA targeting by homopyrimidine PNAs. *Biophys. J.* **72**, 2763-2769.

Egholm, M., Christensen, L., Dueholm, K. L., Buchardt, O., Coull, J. & Nielsen, P. E. (1995). Efficient pH-independent sequence-specific DNA binding by pseudoisocytosine-containing bis-PNA. *Nucl. Acids Res.* **23**, 217-222.

Frank-Kamenetskii, M. D. (1992). Protonated DNA structures. *Methods Enzymol.* **211**, 180-191.

Frank-Kamenetskii, M. D. & Mirkin, S. M. (1995). Triplex DNA structures. *Annu. Rev. Biochem.* **64**, 65-95.

Griffith, M. C., Risen, L. M., Greig, M. J., Lesnik, E. A., Sprankle, K. G., Griffey, R. H., Kiely, J. S. & Freier, S. M. (1995). Single and bis peptide nucleic acids as triplexing agents: binding and stoichiometry. *J. Am. Chem. Soc.* **117**, 831-832.

Howard, F. B., Frazier, J., Lipsett, M. N. & Miles, H. T. (1964). Infrared demonstration of two- and three-strand helix formation between polyC and guanosine mononucleotides and oligonucleotides. *Biochem. Biophys. Res. Commun.* **27**, 93-102.

Izvolsky, K. I., Demidov, V. V., Bukanov, N. O. & Frank-Kamenetskii, M. D. (1998). Yeast artificial chromosome segregation from host chromosomes with similar lengths. *Nucl. Acids Res.* **26**, 5011-5012.

Kuhn, H., Demidov, V. V., Frank-Kamenetskii, M. D. & Nielsen, P. E. (1998). Kinetic sequence discrimination of cationic bis-PNAs upon targeting of double-stranded DNA. *Nucl. Acids Res.* **26**, 582-587.

Kurakin, A., Larsen, H. J. & Nielsen, P. E. (1998). Cooperative strand displacement by peptide nucleic acid (PNA). *Chem. Biol.* **5**, 81-89.

Leitner, D., Schröder, W. & Wiesz, K. (1998). Direct monitoring of cytosine protonation in an intramolecular DNA triple helix. *J. Am. Chem. Soc.* **120**, 7123-7124.

Lomakin, A. & Frank-Kamenetskii, M. D. (1998). A theoretical analysis of nucleic acid interactions with

oligonucleotides and peptide nucleic acids (PNAs). *J. Mol. Biol.* **277**, 57-70.

Lyamichev, V. I., Mirkin, S. M. & Frank-Kamenetskii, M. D. (1985). A pH-dependent structural transition in the homopurine-homopyrimidine tract in superhelical DNA. *J. Biomol. Struct. Dyn.* **3**, 327-338.

Lyamichev, V. I., Mirkin, S. M., Frank-Kamenetskii, M. D. & Cantor, C. R. (1988). A stable complex between homopyrimidine oligomers and the homologous regions of duplex DNAs. *Nucl. Acids Res.* **16**, 2165-2178.

Maher, L. J., Dervan, P. B. & Wold, B. J. (1990). Kinetic analysis of oligodeoxyribonucleotide-directed triple-helix formation on DNA. *Biochemistry*, **29**, 8820-8826.

Møllegaard, N. E., Buchardt, O., Egholm, M. & Nielsen, P. E. (1994). Peptide nucleic acid·DNA strand displacement loops as artificial transcription promoters. *Proc. Natl Acad. Sci. USA*, **91**, 3892-3895.

Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1993). Peptide nucleic acids (PNA): oligonucleotide analogs with a polyamide backbone. In *Antisense Research and Applications* (Crook, S. T. & Lebleu, B., eds), pp. 363-373, CRC Press, Boca Raton.

Nielsen, P. E., Egholm, M. & Buchardt, O. (1994). Evidence for (PNA)₂/DNA triplex structure upon binding of DNA to dsDNA by strand displacement. *J. Mol. Recog.* **7**, 165-170.

Peffer, N. J., Hanvey, J. C., Bisi, J. E., Thomson, S. A., Hassman, F. C., Noble, S. A. & Babiss, L. E. (1993). Strand-invasion of duplex DNA by peptide nucleic acid oligomers. *Proc. Natl Acad. Sci. USA*, **90**, 10648-10652.

Uhlmann, E., Peymann, A., Breipohl, G. & Will, D. W. (1998). PNA: synthetic polyamide nucleic acids with unusual binding properties. *Angew. Chem. Int. Edit.* **37**, 2796-2823.

Veselkov, A. G., Demidov, V. V., Nielsen, P. E. & Frank-Kamenetskii, M. D. (1996a). A new class of genome rare cutters. *Nucl. Acids Res.* **24**, 2483-2488.

Veselkov, A. G., Demidov, V. V., Frank-Kamenetskii, M. D. & Nielsen, P. E. (1996b). PNA as a rare genome-cutter. *Nature*, **379**, 214.

Wilson, W. D., Hopkins, H. P., Mizan, S., Hamilton, D. D. & Zon, G. (1994). Thermodynamics of DNA triplex formation in oligomers with and without cytosine bases: influence of buffer species, pH, and sequence. *J. Am. Chem. Soc.* **116**, 3607-3608.

Wittung, P., Nielsen, P. E. & Norden, B. (1996). Direct observation of strand invasion by peptide nucleic acid (PNA) into double-stranded DNA. *J. Am. Chem. Soc.* **118**, 7049-7054.

Wittung, P., Nielsen, P. E. & Norden, B. (1997). Extended DNA-recognition repetoire of peptide nucleic acid (PNA): PNA-dsDNA triplex formed with cytosine-rich homopyrimidine PNA. *Biochemistry*, **36**, 7973-7979.

*Edited by I. Tinoco*