Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 2637–2641, March 1995
Biochemistry

Kinetics and mechanism of polyamide ("peptide") nucleic acid binding to duplex DNA

VADIM V. DEMIDOV†‡§, MICHAEL V. YAVNILOVICH†¶, BORIS P. BELOTSERKOVSKII†‖, MAXIM D. FRANK-KAMENETSKII††,‡‡, AND PETER E. NIELSEN‡,‡‡

†Institute of Molecular Genetics, Russian Academy of Sciences, Kurchatov Square, 123182 Moscow, Russia; ‡Center for Biomolecular Recognition, Department of Medical Biochemistry and Genetics, Laboratory B, The Panum Institute, Blegdamsvej 3c, DK-2200 Copenhagen N, Denmark; and ††Center for Advanced Biotechnology, Department of Biomedical Engineering, Boston University, 36 Cummington Street, Boston, MA 02215

Communicated by Charles R. Cantor, Boston University, Boston, MA, December 20, 1994

ABSTRACT **To elucidate the mechanism of recognition of double-stranded DNA (dsDNA) by homopyrimidine polyamide ("peptide") nucleic acid (PNA) leading to the strand-displacement, the kinetics of the sequence-specific PNA/DNA binding have been studied. The binding was monitored with time by the gel retardation and nuclease S1 cleavage assays. The experimental kinetic curves obey pseudo-first-order kinetics and the dependence of the pseudo-first-order rate constant, k_{ps}, on PNA concentration, P, obeys a power law $k_{ps} \backsim P^{\gamma}$ with $2 < \gamma < 3$. The k_{ps} values for binding of decamer PNA to dsDNA target sites with one mismatch are hundreds of times slower than for the correct site. A detailed kinetic scheme for PNA/DNA binding is proposed that includes two major steps of the reaction of strand invasion: (*i*) a transient partial opening of the PNA binding site on dsDNA and incorporation of one PNA molecule with the formation of an intermediate PNA/DNA duplex and (*ii*) formation of a very stable PNA_2/DNA triplex. A simple theoretical treatment of the proposed kinetic scheme is performed. The interpretation of our experimental data in the framework of the proposed kinetic scheme leads to the following conclusions. The sequence specificity of the recognition is essentially provided at the "search" step of the process, which consists in the highly reversible transient formation of duplex between one PNA molecule and the complementary strand of duplex DNA while the other DNA strand is displaced. This search step is followed by virtually irreversible "locking" step via PNA_2/DNA triplex formation. The proposed mechanism explains how the binding of homopyrimidine PNA to dsDNA meets two apparently mutually contradictory features: high sequence specificity of binding and remarkable stability of both correct and mismatched PNA/DNA complexes.**

A new type of DNA analogue, polyamide ("peptide") nucleic acid (PNA), was described in 1991 (1). This sequence-specific DNA binding reagent is believed to be a very promising drug (2, 3) with numerous potential applications (4). For homopyrimidine PNAs a unique type of duplex DNA/drug interaction is observed. It consists of PNA binding to one of the DNA strands through formation of stable PNA_2/DNA triplex while the noncomplementary DNA strand is left in single-stranded state (1, 5, 6) thus forming a structure that we call the P loop. P-loop formation leads to selective inhibition of protein binding to DNA (7, 8), results in transcription elongation arrest (2, 7, 9, 10), creates an artificial transcription promoter (11), makes it possible to convert single-strand-specific nucleases into sequence-selective cutters (12), and, if PNA is biotinylated, to place electron-microscopy markers on double-stranded DNA (dsDNA) (13).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

For biomedical and molecular biological applications of PNA it is essential to understand the factors controlling PNA/DNA binding and its sequence specificity. The data indicate that under conditions in which the PNA/DNA complexes are normally studied, the binding is virtually irreversible (5), thus implying a crucial role of kinetic factors in the strand-displacement reaction. Here we present a kinetic study of homopyrimidine PNAs binding to dsDNA and propose a kinetic model for PNA/DNA sequence-specific recognition.

MATERIALS AND METHODS

PNAs H-T_{10}-Lys-NH_2, H-T_5CT_4-Lys-NH_2, and H-$T_2CT_2CT_4$-Lys-NH_2 were synthesized as described (1, 14). The bis-PNA H-T_8-NH_2-R-NH_2-T_8-H consisted of two octathymidyl PNAs coupled in parallel orientation with a flexible N-acetamido-*N,N'*-bis(5-aminopentamethylenecarbonyl)ethylenediamine linker (R).

Target plasmids were constructed by cloning the appropriate oligonucleotides into the polylinker of plasmid pUC19 and were described in detail elsewhere (8, 12). Plasmid pT10 contains the $A_{10}{\cdot}T_{10}$ insert cloned into the *Bam*HI site, pT9C has $A_5GA_4{\cdot}T_4CT_5$ cloned into the *Sal* I site, and pA8G2 has $T_4CT_2CT_2{\cdot}A_2GA_2GA_4$ cloned into the *Pst* I site. ^{32}P-labeled DNA fragments were prepared by cutting the plasmids with restriction enzymes *Eco*RI and *Hin*dIII (or *Pvu* II) and labeling with [α-^{32}P]dNTP by the Klenow fragment of *Escherichia coli* DNA polymerase. PNA binding to dsDNA fragments was performed at 37°C in TE buffer (10 mM Tris·HCl/1 mM EDTA, pH 8.0) containing 10 mM NaCl. Some experiments were performed at higher salt concentrations in a physiological buffer (150 mM KCl/10 mM NaCl/15 mM $MgCl_2$/1 mM $CaCl_2$/1 mM spermidine/10 mM Tris·HCl, pH 6.5). The pH values in concentrated PNA solutions were carefully controlled at 37°C by a micro pH meter. In low-salt kinetic experiments the binding was stopped after the desired times by addition of NaCl up to 100 mM (5) and cooling. Then the samples were analyzed by electrophoresis in 6–10% polyacrylamide gels run in TAE (Tris acetate/EDTA) or TBE (Tris borate/EDTA) buffers. Quantitative analysis (we estimate the accuracy of our experiments to be about 15%) was done by scanning of autoradiographs on a Molecular Dynamics computing densitometer or by direct counting of radioactive bands cut out from dried gels.

Abbreviations: PNA, polyamide ("peptide") nucleic acid; dsDNA, double-stranded DNA.

§Present address: Center for Advanced Biotechnology, Boston University, 36 Cummington Street, Boston, MA 02215.
¶Present address: Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel.
‖Present address: Cell and Molecular Biology Laboratory, SRI International 205-19, 333 Ravenswood Avenue, Menlo Park, CA 94025.
‡‡To whom reprint requests should be addressed.

The time course of PNA binding to dsDNA was quantitatively analyzed in plots of $\ln(D/D_0)$ versus time t, where D is the yield of free DNA fragment (not bound to PNA) and D_0 is the total DNA quantity determined by the sum of the quantities of DNA fragments bound and not bound to PNA.

Pseudo-first-order kinetics was assumed (which proved to be the case; see *Results*):

$$-\ln(D/D_0) = k_{ps}t, \quad [1]$$

where k_{ps} is the pseudo-first-order rate constant. The dependence of k_{ps} on the PNA concentration (P) was approximated as

$$k_{ps} \backsim P^{\gamma}. \quad [2]$$

Thus, the exponent γ was determined as the slope of linear curves in coordinates $\ln \ln(D_0/D)$ versus $\ln P$ for the same period of time in the range of PNA concentrations we worked at. The limitations of the method we used for PNA/DNA complex detection did not allow us to investigate much higher or much lower PNA concentration than we did.

Another method to determine γ was also used. We determined the yield of the fragments, D_{fr}, obtained as a result of digestion of PNA/DNA complexes with nuclease S1 (12). For the initial stage of the kinetics, when the fraction of complexes is small ($D_{fr}/D_0 << 1$),

$$D_{fr}/D_0 \approx k_{ps}t \backsim P^{\gamma}t. \quad [3]$$

Therefore, by plotting $\ln D_{fr}$ versus $\ln P$ for the same initial DNA quantities and the same periods of time, one can estimate γ.

RESULTS

To study the kinetics of the binding of homopyrimidine PNAs to dsDNA, we at first used a gel retardation assay. The radiolabeled DNA fragments were complexed with PNA and the PNA/DNA complex was separated from free DNA by electrophoresis in a polyacrylamide gel with subsequent autoradiographic quantification of the intensities of both bands. Fig. 1*A* shows a typical pattern. The intensity of the faster migrating band gives the quantity of free DNA, whereas the slower migrating band corresponds to the complex, which is retarded due to P-loop formation (7–9). By measuring the intensities of these two bands, normalized with their integral intensity, one can follow the kinetics of the PNA binding to DNA.

Fig. 1*B* shows the kinetic data on the binding of 10 μM PNA H-T_{10}-Lys-NH_2 with the perfectly matched (A_{10}·T_{10}) and mismatched (A_5GA_4·T_4CT_5) sites on the corresponding linear DNA fragments. Fig. 1*C* demonstrates that the binding of homopyrimidine PNAs to DNA sites follows the pseudo-first-order kinetics described by Eq. **1**. One observes a dramatic difference in the values of the pseudo-first-order rate constant, k_{ps}, for the perfect and the mismatched sites. With PNA H-T_{10}-Lys-NH_2, $k_{ps} = 4.2 \times 10^{-4}$ min^{-1} for the mismatched site, 100 times lower than for the perfect one ($k_{ps} = 3.6 \times 10^{-2}$ min^{-1}).

With PNA H-T_5CT_4-Lys-NH_2 this difference is even greater: we could not detect any binding of this PNA at 10 μM for 3 days of incubation with plasmid pT10, carrying a single-mismatch insert, A_{10}·T_{10}, while binding to pT9C, carrying the perfect target, A_5GA_4·T_4CT_5, was completed within an hour. This corresponds to a >1000-fold decrease of the rate of binding to the mismatched target as compared with the rate of binding to the perfect target. Only upon increasing the H-T_5CT_4-Lys-NH_2 concentration by a factor of 10 did we observe binding of this PNA to pT10, with $k_{ps} = 5 \times 10^{-3}$ min^{-1} (Fig. 1*C*). Proceeding from the value $k_{ps} = 0.11$ min^{-1} for the binding of this PNA at 10 μM to the true target (Fig. 1*C*) and taking the nonlinear, approximately quadratic, dependence of k_{ps} on PNA concentration (see below) into account, we can estimate that for the binding of 100 μM H-T_5CT_4-Lys-NH_2 to the true site, $k_{ps} = 10$ min^{-1}, a value 2000 times higher than for the mismatched site. This estimation of the sequence selectivity of PNA binding is very close to the one arrived at above. Note that in spite of a very low rate of PNA binding to mismatched DNA sites, complete PNA/DNA binding can be obtained by incubating for a long time at a high PNA concentration. Thus these mismatched complexes are sufficiently thermodynamically stable.

By employing gel retardation analysis as in Fig. 1*A* and varying the PNA concentration at a fixed incubation time, one can determine the exponent γ in Eq. **2** by plotting the results in $\ln \ln(D_0/D)$ versus $\ln P$ coordinates. The experimental points lie on a straight line (Fig. 2), indicating that the power approximation of Eq. **2** is valid, at least within the range of PNA concentrations studied, yielding the exponent values γ =

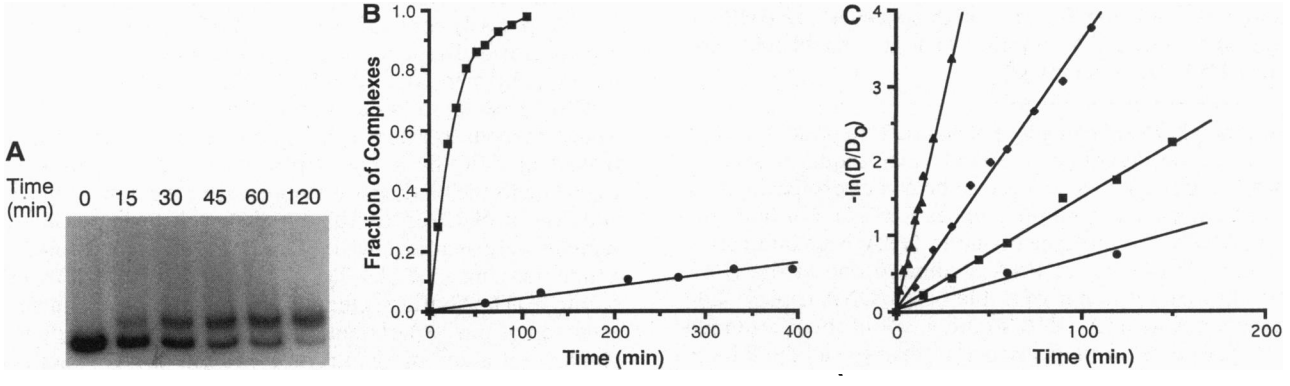


FIG. 1. PNA/DNA binding kinetics. (*A*) Autoradiograph of the 10% polyacrylamide gel after electrophoretic separation of the pure DNA fragment and the fragment complexed with PNA after incubation for the indicated times. The ^{32}P-labeled *Eco*RI–*Pvu* II fragment of plasmid pA8G2 was incubated with 25 μM PNA H-$T_2CT_2CT_4$-Lys-NH_2 in 0.5× TE buffer (pH 8.0) with 10 mM NaCl at 37°C. (*B*) Fraction of DNA fragments carrying inserts A_{10}·T_{10} (■) or A_5GA_4·T_4CT_5 (●) complexed with PNA H-T_{10}-Lys-NH_2 as a function of the incubation time. Radiolabeled *Eco*RI–*Hin*dIII fragments of pT10 and pT9C were incubated for various times with 10 μM H-T_{10}-Lys-NH_2 in TE buffer containing 10 mM NaCl at 37°C. (*C*) Semilogarithmic plot of the kinetic data on PNA binding with the true and mismatched DNA targets, where D/D_0 is the fraction of DNA fragments not bound to PNA. The plasmid pT10 fragment was incubated with 10 μM H-T_{10}-Lys-NH_2 (◆) in TE buffer or with 100 μM PNA H-T_5CT_4-Lys-NH_2 (●) in 0.5× TE buffer, the plasmid pT9C fragment was incubated with 10 μM PNA H-T_5CT_4-LysNH_2 (▲) in 0.5× TE buffer, and the plasmid pA8G2 fragment was incubated with 25 μM PNA H-$T_2CT_2CT_4$-Lys-NH_2 (■) in 0.5× TE buffer, all at 37°C and pH 8.0 in 10 mM NaCl. The values of the rate constant are 11, 3.6 and 1.5 × 10^{-2} min^{-1} for correct binding of PNAs H-T_5CT_4-Lys-NH_2, H-T_{10}-Lys-NH_2, and H-$T_2CT_2CT_4$-Lys-NH_2, respectively.

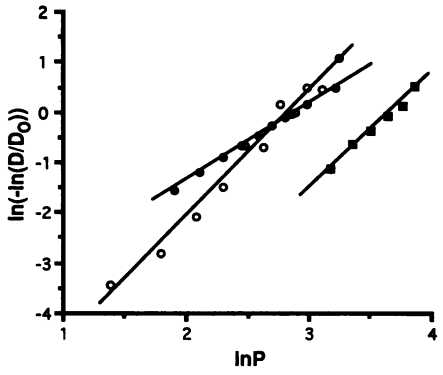


FIG. 2. Double logarithmic plot of the fraction D/D_0 of plasmid DNAs, carrying the appropriate inserts, not bound to the PNAs as a function of PNA concentration (μM) at 37°C in TE buffer at pH 8.0 in 10 mM NaCl. The plasmid pT10 fragment was incubated for 30 min with "monomeric" PNA H-T_{10}-Lys-NH_2 (○) or with "dimeric" PNA (H-T_8-NH_2-R-NH_2-T_8-H) (●) and the plasmid pA8G2 fragment was incubated for 15 min with PNA H-$T_2CT_2CT_4$-Lys-NH_2 (■).

2.5 and 2.2 for H-T_{10}-Lys-NH_2 and H-$T_2CT_2CT_4$-Lys-NH_2, respectively. In a similar experiment for a bis PNA consisting of two PNA T_8 moieties connected by a flexible linker, $\gamma = 1.5$, lower by 1 unit than in the case of the "monomeric" PNA H-T_{10}-Lys-NH_2 (Fig. 2).

We also used nuclease S1 probing of the strand-displaced PNA/DNA complex (12) to follow the binding kinetics by analyzing the yield of corresponding fragments in agarose gels. We used this assay to study the concentration dependence of binding of PNA H-T_{10}-Lys-NH_2 to its perfect binding site A_{10}·T_{10} at two Na^+ concentrations. From the initial part of the binding kinetics the exponent γ in Eq. **2** can be determined as the slope of plots of ln D_{fr} versus ln P. These data yield the value $\gamma = 2.7$ (Fig. 3), which agrees well with the γ value determined by the gel retardation assay. Thus, the PNA concentration dependences of k_{ps} for all PNAs studied are nonlinear, and for H-T_{10}-Lys-NH_2 the dependence is close to cubic, at least within the concentration range we studied. The rate of PNA binding to dsDNA decreases with the increase in salt concentration, in agreement with earlier observations (5, 9).

We conclude that at low salt and the PNA concentrations used in our experiments the equilibrium is strongly shifted toward complex formation but that strand invasion can require long incubation because of slow kinetics. At high salt the situation appears qualitatively similar, although much higher PNA concentrations are required to observe formation of the PNA/DNA complex: 50% binding of PNA H-$T_2CT_2CT_4$-Lys-NH_2 to a complementary target of plasmid pA8G2 in physiological buffer, containing all major cellular mono- and polycations, occurs within 1 hr at a PNA concentration of 350 μM, and increasing the PNA up to 500 μM leads to complete binding during this period of time. Under the same conditions this PNA does not bind detectably to plasmid pT9C plasmid, containing a target with one mismatch.

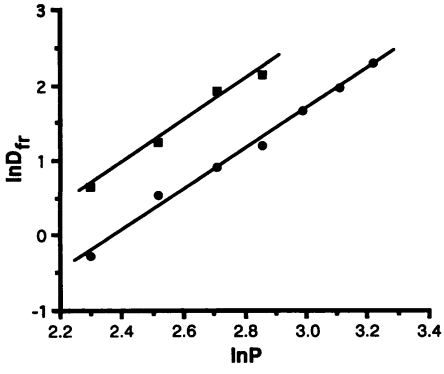


FIG. 3. Nuclease S1 digestion of linearized pT10 at various concentrations of PNA H-T_{10}-Lys-NH_2. The PNA concentration was varied from 10 to 25 μM for 20 min incubation of plasmid DNA with PNA at 10 mM (■) or 20 mM (●) NaCl. D_{fr} for binding of 25 μM PNA was taken as 10 relative units.

KINETIC MODEL

All data available (4) are fully consistent with the PNA/DNA binding mode in which two PNAs are required for the formation of a stable strand-displacement complex involving both Watson–Crick and Hoogsteen base pairing (15). Several plausible, closely related models for PNA binding to dsDNA, all involving the inherent "breathing" of the base pairs in the DNA double helix, were considered (2, 4, 6, 8, 16). Fig. 4*A* shows the model and kinetic scheme we use here to analyze our data. First, a PNA molecule binds to a fluctuationally open part of the binding site via Watson–Crick complementary

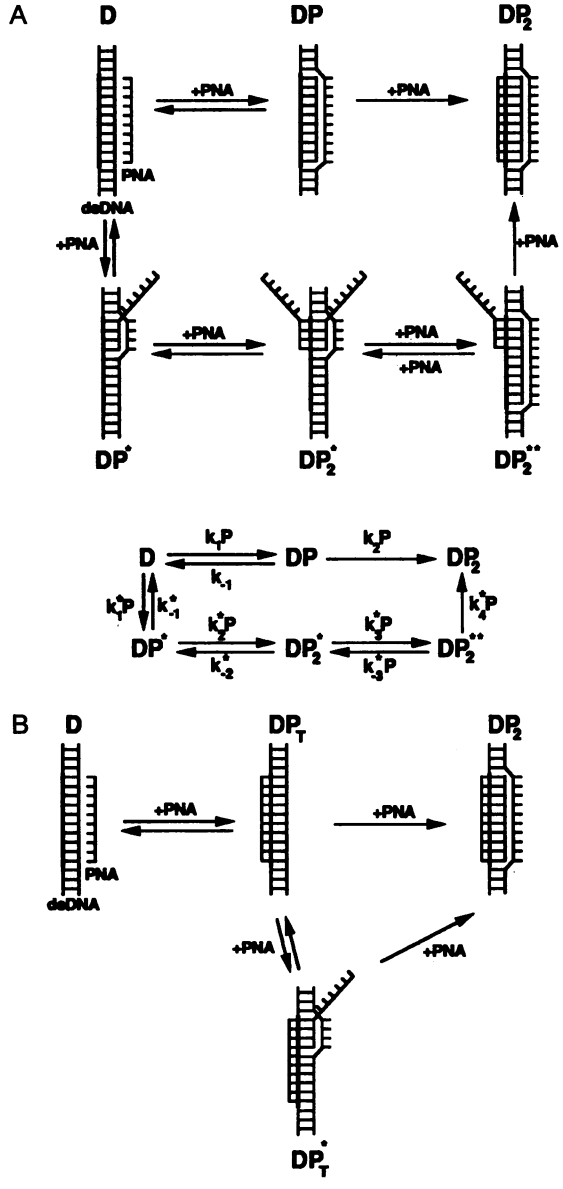


FIG. 4. Schematics showing possible pathways of PNA/DNA strand-displacement reaction. (*A*) The pathways (*Upper*) and the corresponding scheme (*Lower*) used for kinetic analysis. (*B*) Alternative scheme for PNA/DNA binding involving a PNA/dsDNA triplex intermediate.

pairing. For simplicity, we consider only two types of these intermediate complexes: one, designated DP, with a correctly positioned full-length PNA/DNA duplex and one, DP*, with a partial duplex, in which the PNA molecule is incorrectly positioned with respect to its binding site. We can illustrate two possible examples of such "shifted" complexes for DNA targets in plasmids pT10 and pA8G2.

```
          TTTTTTTTTT   TTTTCTTCTT
          ••••••       ••••
5'-ccAAAAAAAAAAgg-3'       5'-agAAGAAGAAAAcc-3'
```

The sequences of PNA-T_{10} and PNA-$T_2CT_2CT_4$ are shown in bold type and DNA target sequences are capitalized. In practice, these partial duplexes are numerous. All of the above intermediates, however, are very unstable. We assume that only after the DP duplex is trapped by the second PNA molecule, forming a triplex, is the very stable, virtually irreversible, complex DP_2 formed. The "shifted" intermediate DP* can also form the more stable transient complex (DP_2^*) as a result of addition of the second PNA strand. We assume that once such a shifted triplex is formed it intensifies "breathings" of the adjacent DNA regions, thus greatly enhancing accessibility of the parts of the DNA site which remain in the duplex state to the next PNA molecule. This third PNA molecule forms the correct duplex with the same site, pushing out the shifted duplex, thus making intermediate DP_2^{**}. The effect of enhanced accessibility of the regions adjacent to the DNA open site is well known from kinetic studies of the reaction of DNA with formaldehyde (17). Finally, the transient DP_2^{**} complex is also converted into the stable DP_2 triplex when yet another PNA molecule pushes out the shifted second PNA strand in the triplex.

Note that we consider the final stage of the formation of the "correct" triplex as irreversible, whereas we consider the formation of "incorrect" triplexes (like DP_2^* species) as reversible. Of course, the final stage is also reversible but the reverse kinetic constant is very small when the conditions favor PNA_2/DNA triplex formation. Under such conditions the kinetic constant of dissociation of the complex is known to depend exponentially on the length of the complex (18, 19). This explains why the "incorrect" triplexes are predominantly short-lived whereas the "correct" triplex is very long-lived. We also do not consider for the first approximation all possible transitions between shifted and correct intermediates, because this greatly simplifies the calculations and evidently misses nothing of principal importance.

Thus, let us designate by D the concentration of free DNA; by P the concentration of free PNA; by DP and DP^* the concentrations of complete and incomplete DNA complexes with one PNA molecule, respectively; by DP_2^* and DP_2^{**} the concentrations of transitional complexes with two PNA molecules; and by DP_2 the concentration of the resulting PNA_2/DNA triple complex. The following kinetic equations can be derived from the kinetic scheme presented in Fig. 4*A*:

$$dD/dt = -(k_1 + k_1^*)(D)(P) + k_{-1}DP + k_{-1}^*DP^* \quad [4]$$

$$dDP/dt = k_1(D)(P) - k_{-1}DP - k_2(DP)(P) \quad [5]$$

$$dDP^*/dt = k_1^*(D)(P) + k_{-2}^*DP_2^* - k_{-1}^*DP^* - k_2^*(DP^*)(P) \quad [6]$$

$$dDP_2^*/dt = k_2^*(DP^*)(P) + k_{-3}^*(DP_2^{**})(P) - k_{-2}^*DP_2^* - k_3^*(DP_2^*)(P) \quad [7]$$

$$dDP_2^{**}/dt = k_3^*(DP_2^*)(P) - (k_{-3}^* + k_4^*)(DP_2^{**})(P) \quad [8]$$

$$dDP_2/dt = k_2(DP)(P) + k_4^*(DP_2^{**})(P). \quad [9]$$

In practice, one needs PNA at $\approx$10 μM to observe the strand-displacement reaction within $\approx$1 hr. At the same time, the usual DNA concentration—i.e., the concentration of the PNA binding sites—is $\approx$0.05 μM. This means that the number of PNA molecules in our kinetic experiments exceeds the number of DNA molecules by a factor of 200. Therefore, the free PNA concentration does not decrease during the reaction—i.e., P is a constant. As indicated above, all intermediates are so unstable that we can assume that their concentration is low and does not change significantly during the reaction—i.e., $dDP/dt = 0$, $dDP^*/dt = 0$, $dDP_2^*/dt = 0$, and $dDP_2^{**}/dt = 0$. From Eqs. **4–9**, under the above quasistationary assumptions and for the initial conditions $D(0) = D_0$, $DP_2(0) = 0$, we obtain

$$D = D_0\exp(-k_{ps}t) \quad [10]$$

$$DP_2 = D_0[1 - \exp(-k_{ps}t)] \quad [11]$$

$$k_{ps} = K_1k_2P^2 + K_1^*K_2^*K_3^*k_4^*P^3, \quad [12]$$

where

$$K_1 = k_1/(k_{-1} + k_2P) \quad [13]$$

$$K_1^* = k_1^*/[k_{-1}^* + (k_2^* - k_{-2}^*K_2^*)P] \quad [14]$$

$$K_2^* = k_2^*/[k_{-2}^* + k_3^*k_4^*\,P/(k_{-3}^* + k_4^*)] \quad [15]$$

$$K_3^* = k_3^*/(k_{-3}^* + k_4^*). \quad [16]$$

Thus, we arrive at the conclusion that under the above assumption our complicated many-step kinetic scheme obeys the pseudo-first-order kinetics (Eqs. **10** and **11**) with complex nonlinear dependence of the pseudo-first-order rate constant on the PNA concentration, P (Eqs. **12–16**). If we assume that the transient complexes are very unstable so that the kinetic constants and the PNA concentration correspond to the case of $k_{-1} \gg k_2P$, $k_{-3}^* \gg k_4^*$, and $k_{-2}^* \gg k_3^*\,k_4^*P/k_{-3}^*$, then all K_i values become independent of P ($K_1 = k_1/k_{-1}$, $K_1^* = k_1^*/k_{-1}^*$, $K_2^* = k_2^*/k_{-2}^*$, $K_3^* = k_3^*/k_{-3}^*$), and Eq. **12** is simplified to yield a simple polynomial dependence of k_{ps} on P as a sum of quadratic and cubic terms. This results in a curve of ln k_{ps} versus ln P with the slope varying from 2 to 3. The latter means that for this case the dependence of the pseudo-first-order rate constant on P can be approximated, at least in a limited range of variation of P, as the polynomial function 1 (Eq. **2**) with $2 < \gamma < 3$. The γ values we observe in our experiments lie within these limits. The fact that $\gamma > 2$ indicates that the second term in Eq. **12** is significant, which indicates that the unstable "shifted" intermediates play an important role in the overall kinetics. It is tempting to interpret the fact that γ is larger for the regular PNA H-T_{10}-Lys-NH_2 than for the less regular PNA H-$T_2CT_2CT_4$-Lys-NH_2 in terms of a larger role of the "shifted" intermediates in the former case. More accurate measurements of the γ values for PNAs with different sequences are necessary to establish such a correlation between the γ value and regularity of the sequence.

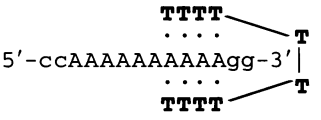
In the case of bis PNAs, the final triplex is formed by clamping of bis PNA and its formation does not depend on the PNA concentration. As a result, we have to substitute the products k_2P and k_4^*P in Eqs. **12**, **13**, and **15** for k_2 and k_4^*, respectively, to obtain the modified equation for k_{ps}^{bis}.

$$k_{ps}^{bis} = K_1k_2P + K_1^*K_2^*K_3^*k_4^*P^2, \quad [17]$$

where K_1 and K_2^* are described by the modified Eq. **13** and Eq. **15**, respectively. Therefore, for bis PNA the exponent γ in Eq. **2** should be 1 less than in the case of monomeric PNA, which agrees perfectly with our experimental findings. Again, the fact that in this case exponent $\gamma > 1$ indicates that for bis PNA the shifted DP^* intermediates also play an important role at the initial step of binding.

Our kinetic scheme is the simplest one that explains the major features of our experimental data. Other intermediates

are possible that have not been included in our model at the present stage. One of the alternative intermediates is the "self-clamped" PNA complex with the part of the DNA target site.



This intermediate looks very attractive because it can form in the correct register (if it has the opposite orientation in the above example). This complex could be readily broken up by the second PNA molecule with the formation of the correct PNA_2/DNA triplex. However, such self-clamped intermediates can be formed only by fully or partially symmetric PNAs.

Another possible intermediate is the earlier proposed PNA/dsDNA triplex (2, 4, 16). The formal plausible scheme of interconversion with the participation of this intermediate (Fig. 4*B*) shows that the value of γ in this case may also be >2. The PNA strand in the PNA/dsDNA triplex intermediate may "swap" hydrogen-bonding positions, going from Hoogsteen to Watson–Crick, thereby displacing the pyrimidine DNA strand. Thus, this alternative scheme may have some common features with the scheme discussed above, and more thorough kinetic analysis is required to distinguish between these or take both into account for some special cases of PNA/dsDNA interaction.

DISCUSSION

Under our experimental conditions the sequence specificity of PNA/DNA recognition is essentially determined by the equilibrium constant of formation of the Watson–Crick duplex between PNA and the complementary strand of the DNA target site. Indeed, the above analysis indicates that our experimental situation corresponds to the limiting case of $k_{-1} \gg k_2P$, $k^*_{-3} \gg k^*_4$, $k^*_{-2} \gg k^*_3k^*_4P/k^*_{-3}$. Eqs. **12–16** show that in this case the kinetic constant of the overall process, k_{ps}, is proportional to the equilibrium constants for the reactions of DNA/PNA duplex formation.

One should therefore expect the site specificity of recognition (i.e., the selectivity of k_{ps} with respect to different sequences) will depend on the free-energy difference for the mismatched duplex and the regular duplex. The shift in melting temperatures due to different mismatches is in the same range for DNA/DNA, DNA/PNA, and PNA/PNA duplexes (20, 21), and for the DNA/DNA case the free-energy loss due to mismatches was estimated to be 3–4 kcal/mol (22). Therefore, k_{ps} should be 2–3 orders of magnitude smaller for a site with one mismatch than for the correct site. This estimation nicely agrees with our experimental observations (see Fig. 1 *B* and *C*). The two situations of mismatched PNA/DNA binding we studied correspond to two extreme values of a mismatch's energy cost. The mismatched pair PNA-T·DNA-G corresponds to the minimal energy cost, whereas the mismatched pair PNA-C·DNA-A corresponds to the maximal energy cost (20, 23). Thus, the total sequence selectivity of k_{ps} we observe can be entirely attributed to the selectivity of the formation of the transient DNA/PNA duplexes. This indicates that the final, irreversible step of our kinetic scheme in Fig. 4, PNA_2/DNA triplex formation, is much less sequence selective, at least under our experimental conditions.

CONCLUSION

The data and their theoretical analysis show that sequence-specific recognition of duplex DNA by homopyrimidine PNA is determined, at least under our experimental conditions, by equilibrium formation of the duplex between one PNA molecule and the complementary strand of DNA via a transient strand-displacement reaction. This intermediate complex, however, is highly unstable and requires either a high concentration of PNA at physiological salt concentration or very low salt. Only after this intermediate is eventually converted into a PNA_2/DNA triplex is a very stable, virtually irreversible PNA/dsDNA complex formed. The proposed mechanism explains how PNA binding to DNA matches two apparently mutually contradictory features: high sequence specificity of binding with remarkable stability of both correct and singly mismatched PNA/DNA complexes.

This paper is dedicated to the memory of the late Professor Ole Buchardt. We thank Professors D. M. Crothers and Yu. S. Lazurkin for very useful discussions. We appreciate the technical help of Dr. N. Bukanov in the preparation of the manuscript. This work was supported by ISIS Pharmaceuticals (Carlsbad, CA) and by Grant M4J000 from the International Science Foundation and Grant 11424-a from the Russian Foundation for Basic Research.

1. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1991) *Science* **254,** 1497–1500.
2. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1993) *Anti-Cancer Drug Des.* **8,** 53–63.
3. Demidov, V. V., Potaman, V. N., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., Sönnichsen, S. H. & Nielsen, P. E. (1994) *Biochem. Pharmacol.* **48,** 1310–1313.
4. Nielsen, P. E., Egholm, M. & Buchardt, O. (1994) *Bioconj. Chem.* **5,** 3–7.
5. Cherny, D. Y., Belotserkovskii, B. P., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O., Berg, R. H. & Nielsen, P. E. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 1667–1670.
6. Almarsson, Ö., Bruice, T. C., Kerr, J. & Zuckermann, R. N. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7518–7522.
7. Hanvey, J. C., Peffer, N. J., Bisi, J. E., Thomson, S. A., Cadilla, R., Josey, J. A., Ricca, D. J., Hassman, C. F., Bonham, M. A., Au, K. G., Carter, S. G., Bruckenstein, D. A., Boyd, A. L., Noble, S. A. & Babiss, L. E. (1992) *Science* **258,** 1481–1485.
8. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1993) *Nucleic Acids Res.* **21,** 197–200.
9. Peffer, N. J., Hanvey, J. C., Bisi, J. E., Thomson, S. A., Hassman, C. F., Noble, S. A. & Babiss, L. E. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 10648–10652.
10. Nielsen, P. E., Egholm, M. & Buchardt, O. (1994) *Gene* **149,** 139–145.
11. Møllegaard, N. E., Buchardt, O., Egholm, M. & Nielsen, P. E. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 3892–3895.
12. Demidov, V., Frank-Kamenetskii, M. D., Egholm, M., Buchardt, O. & Nielsen, P. E. (1993) *Nucleic Acids Res.* **21,** 2103–2107.
13. 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) *Nucleic Acids Res.* **22,** 5218–5222.
14. Egholm, M., Buchardt, O., Nielsen, P. E. & Berg, R. H. (1992) *J. Am. Chem. Soc.* **114,** 9677–9678.
15. Nielsen, P. E., Egholm, M. & Buchardt, O. (1994) *J. Mol. Recognit.* **7,** 165–170.
16. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. (1993) in *Antisense Research and Application*, eds. Crook, S. & Lebleu, B. (CRC, Boca Raton, FL), pp. 363–373.
17. Frank-Kamenetskii, M. D. (1985) in *Structure and Motion: Membranes, Nucleic Acids and Proteins*, eds. Clementi, E., Corongiu, G., Sarma, M. H. & Sarma, R. H. (Adenine, New York), pp. 417–432.
18. Craig, M. E., Crothers, D. M. & Doty, F. (1971) *J. Mol. Biol.* **62,** 383–401.
19. Anshelevich, V. V., Vologodskii, A. V., Lukashin, A. V. & Frank-Kamenetskii, M. D. (1984) *Biopolymers* **23,** 39–58.
20. Egholm, M., Buchardt, O., Christensen, L., Behrens, C., Freier, S. M., Driver, D. A., Berg, R. H., Kim, S. K., Norden, B. & Nielsen, P. E. (1993) *Nature (London)* **365,** 566–568.
21. Wittung, P., Nielsen, P. E., Buchardt, O., Egholm, M. & Norden, B. (1994) *Nature (London)* **368,** 561–563.
22. Tibanyenda, N., de Bruin, S. H., Haasnoot, C. A. G., van der Marel, G. A., van Boom, J. H. & Hilbers, C. W. (1984) *Eur. J. Biochem.* **139,** 19–27.
23. Egholm, M., Behrens, C., Christensen, L., Berg, R. H., Nielsen, P. E. & Buchardt, O. (1993) *J. Chem. Soc. Chem. Commun.* **9,** 800–801.