Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 3806–3810, May 1993
Biochemistry

Sequence-specific recognition of double helical RNA and RNA·DNA by triple helix formation

HOGYU HAN AND PETER B. DERVAN

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125

Contributed by Peter B. Dervan, January 26, 1993

ABSTRACT **The stabilities of eight triple helical pyrimidine·purine·pyrimidine structures comprised of identical sequence but different RNA (R) or DNA (D) strand combinations were measured by quantitative affinity cleavage titration. The differences in equilibrium binding affinities reveal the importance of strand composition. For the sequences studied here, the stabilities of complexes containing a pyrimidine third strand D or R and purine·pyrimidine double helical DD, DR, RD, and RR decrease in order: D + DD, R + DD, R + DR, D + DR > R + RD, R + RR >> D + RR, D + RD (pH 7.0, 25°C, 100 mM NaCl/1 mM spermine). These findings suggest that RNA and DNA oligonucleotides will be useful for targeting (*i*) double helical DNA and (*ii*) RNA·DNA hybrids if the purine Watson–Crick strand is DNA. However, RNA, but not DNA, oligonucleotides will be useful for sequence-specific binding of (*i*) double helical RNA and (*ii*) RNA·DNA hybrids if the purine Watson–Crick strand is RNA. This has implications for the design of artificial ligands targeted to specific sequences of double helical RNA and RNA·DNA hybrids.**

Oligodeoxyribonucleotide-directed triple helix formation is one of the most versatile methods for the sequence-specific recognition of double helical DNA (1, 2). The ability to target a broad range of DNA sequences, the stabilities of deoxyribonucleic acid triple helical complexes, and the sensitivity to single base mismatches allow single site targeting within megabase DNA (3–5). At least two structural classes of DNA triple helices exist that differ in base triplet interactions, sequence compositions of the third strand, relative orientations, and relative positions of the phosphate-deoxyribose backbones. Pyrimidine oligodeoxyribonucleotides bind purine tracts in the major groove of double helical DNA parallel to the purine Watson–Crick strand (1). Sequence specificity is derived from thymine (T) recognition of adenine·thymine (AT) base pairs (T·AT base triplets) and N-3 protonated cytosine (C^+) recognition of guanine·cytosine (GC) base pairs (C+GC base triplets) (6, 7). Purine-rich oligodeoxyribonucleotides bind purine tracts in the major groove of DNA antiparallel to the Watson–Crick purine strand (8, 9). Sequence specificity is derived from guanine (G) recognition of GC base pairs (G·GC base triplets) and from A or T recognition of AT base pairs (A·AT and T·AT triplets) (8–10). The stabilities of triple helical complexes are dependent on the length, sequence composition, and temperature, as well as solution conditions, including pH and cation concentrations (11–13). Dissecting the relative contributions of all factors controlling triple helix stabilities will be pivotal when considering the use of oligonucleotide-directed sequence-specific recognition of double helical nucleic acids for *in vivo* applications where temperature, pH, and salt conditions are strictly controlled.

For sequence-specific recognition of double helical RNA or DNA·RNA hybrids by triple helix formation, the importance of strand composition on the stabilities of triple helical complexes must be determined. In a formal sense, there are eight triple helical complexes wherein each strand could consist of DNA or RNA (see Fig. 1). We report the thermodynamic stabilities of eight triple helical complexes measured by quantitative affinity cleavage titration at pH 7.0 and 25°C. The results reveal that triple helical stabilities are different depending on backbone composition. Although in qualitative agreement with stabilities reported recently by Roberts and Crothers (14), there are differences in order and scale.

MATERIALS AND METHODS

General. Sonicated, deproteinized calf thymus DNA (Pharmacia) was dissolved in H_2O to a final concentration of 2 mM in base pairs (bp). Ribonucleotide triphosphates were Pharmacia ultrapure grade and were used as supplied. [γ-^{32}P]ATP (≥3000 Ci/mmol; 1 Ci = 37 GBq) was obtained from Amersham. Calf intestine alkaline phosphatase was purchased from Pharmacia and T4 polynucleotide kinase was obtained from New England Biolabs. Phosphoramidites were purchased from Applied Biosystems (dA, dG, dC, and T) and from BioGenex Laboratories (San Ramon, CA) (rA, rG, rC, U).

Oligodeoxyribonucleotide-EDTA and Oligoribonucleotide-EDTA Syntheses. Pyrimidine oligodeoxyribonucleotides (D) and oligoribonucleotides (R) were synthesized by standard automated solid-support chemistry using an Applied Biosystems model 394 DNA/RNA synthesizer and cyanoethyl *N*,*N*-diisopropyl phosphoramidites for DNA and 2′-*t*-butyldimethylsilyl 3′-cyanoethyl *N*,*N*-diisopropyl phosphoramidites for RNA. The synthesis of oligonucleotides D and R containing EDTA at the 5′ end (*T and *u) will be described elsewhere (unpublished data).

Double Helical DNA and RNA Preparation. DNA templates, 36 nt in length, were prepared by chemical synthesis. RNA templates, 36 nt in length, were prepared by enzymatic synthesis using T7 RNA polymerase (Pharmacia) employing a 53-nt single-stranded DNA template (15). DNA and RNA were purified with 15% denaturing polyacrylamide gels and extracted with 0.3 M NaCl or NaOAc, respectively. DNA was 5′-end labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase, precipitated with ethanol, and dissolved in H_2O (16). RNA was dephosphorylated with calf intestine alkaline phosphatase and subsequently 5′-^{32}P labeled with T4 polynucleotide kinase, precipitated, and dissolved in H_2O. The 5′-^{32}P-end-labeled duplex was prepared by mixing labeled strands with their unlabeled complementary strands in hybridization buffer (200 mM NaCl/1 mM EDTA) at 37°C and purified with a 15% nondenaturing polyacrylamide gel. The 5′-end-labeled duplex was eluted with 200 mM NaCl and filtered through a 0.45-μm filter, followed by ethanol precipitation.

Quantitative Affinity Cleavage Titrations. Cleavage titration experiments were performed in association buffer [50 mM Tris acetate, pH 7.0/100 mM NaCl/1 mM spermine/200

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.

Abbreviations: R, RNA strand; D, DNA strand.

μM (in terms of bp) calf thymus DNA] containing 20,000 cpm of duplex. The DNA and the various concentrations of oligonucleotide-EDTA·Fe were allowed to equilibrate for 24 hr at 25°C. The cleavage reactions were initiated by addition of dithiothreitol (4 mM) and incubated for 8 hr at 25°C. The cleavage reactions were stopped by ethanol precipitation and analyzed by gel electrophoresis.

Quantitation. Gels were exposed to photostimulable storage phosphor imaging plates (Kodak storage phosphor screen S0230 obtained from Molecular Dynamics) in the dark at 25°C for 24 hr. A Molecular Dynamics 400S PhosphorImager was used to obtain data from the phosphorimaging screens. Rectangles of the same dimensions were drawn around the cleavage bands at the target and at the reference sites. The IMAGEQUANT version 3.0 program running on an AST Premium 386/33 computer was used to integrate the volume of each rectangle.

Affinity Cleavage Titration Fitting Procedure. A detailed description of the affinity cleavage titration procedure used has been published (11, 12). The relative cleavage efficiencies at the target site for each oligonucleotide concentration were determined by using the following equation:

$$I_{site} = I_{tot} - \lambda I_{ref}, \quad [1]$$

where I_{tot} and I_{ref} are the intensities of the cleavage bands at the target site and at the reference site, respectively, and λ is a scaling parameter defined as the ratio I_{tot}/I_{ref} at the lowest oligonucleotide concentration. A binding curve, represented by Eq. 2, where I_{sat} is the apparent maximum cleavage, K_T is the equilibrium association constant for the oligonucleotide, and $[O]_{tot}$ is the oligonucleotide-EDTA concentration, was used to fit the experimental data using I_{sat} and K_T as adjustable parameters:

$$I_{site} = I_{sat} \cdot \frac{K_T[O]_{tot}}{1 + K_T[O]_{tot}}. \quad [2]$$

KALEIDAGRAPH software (version 2.1, Abelweck Software; Synergy Software, Reading, PA) running on a Macintosh IIfx computer was used to minimize the difference between I_{fit} and I_{site} for all data points. All reported values are the means of three experimental observations ± SEM. For graphical representation and comparison, I_{site} values were converted to fractional occupancies by dividing I_{site} by I_{sat}.

RESULTS

Affinity Cleavage. Two pyrimidine strands, 18 nt in length, consisting of DNA (D) and RNA (R) were synthesized with EDTA on the T or U base at the 5′ end for affinity cleavage reactions. Four 36-nt strands were synthesized to form stable 35-bp duplexes, each containing an 18-bp site for triple helix formation: all DNA (DD), DNA purine strand + RNA pyrimidine strand (DR), RNA purine strand + DNA pyrimidine strand (RD), and all RNA (RR). Affinity cleavage reactions were performed on the eight triple helical complexes formed upon binding of the 18-nt D- or R-EDTA·Fe conjugates to each of the four duplex molecules (Figs. 1 and 2). To avoid any complication due to differences in oxidative cleavage of DNA and RNA, both strands of DD, DR, RD, and RR were labeled in separate experiments. Each of the eight binding affinities was examined independently in two separate experiments. Only six of eight possible triple helical complexes were formed. At 2 μM concentration of D and R (pH 7.0, 25°C), oligonucleotide D forms a stable triple helix only with DD and DR duplexes, whereas oligonucleotide R forms a stable complex with all four duplexes DD, DR, RD, and RR. For all cases where the triple helical complex forms, oligonucleotides D and R bind parallel to the purine strand of the Watson–Crick duplexes, DD, DR, RD, and RR.

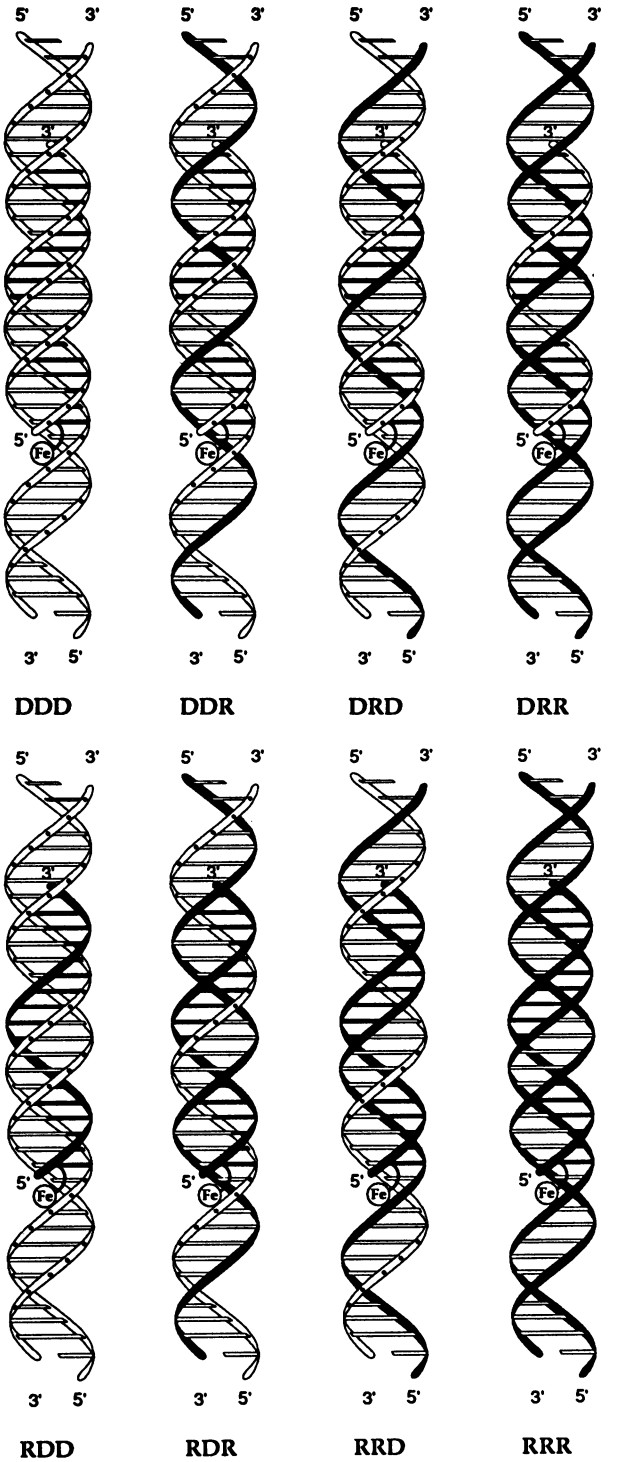

FIG. 1. Ribbon models of eight triple helical pyrimidine·purine·pyrimidine structures formed by binding of an 18-nt pyrimidine oligonucleotide-EDTA·Fe to the 18-bp target purine sequence within the 35-bp double helix. The DNA and RNA strands are depicted as white and dark ribbons, respectively. The Hoogsteen pyrimidine strand is indicated by the first letter, and the Watson–Crick purine and pyrimidine strands are indicated by the second and third letters, respectively. The conformations of these eight triple helical structures are unknown and, therefore, are uniformly represented here.

Quantitative Affinity Cleavage Titration. The association constants for RNA and DNA oligonucleotides-EDTA·Fe D and R (18 nt) binding to 18-bp purine·pyrimidine tracts within

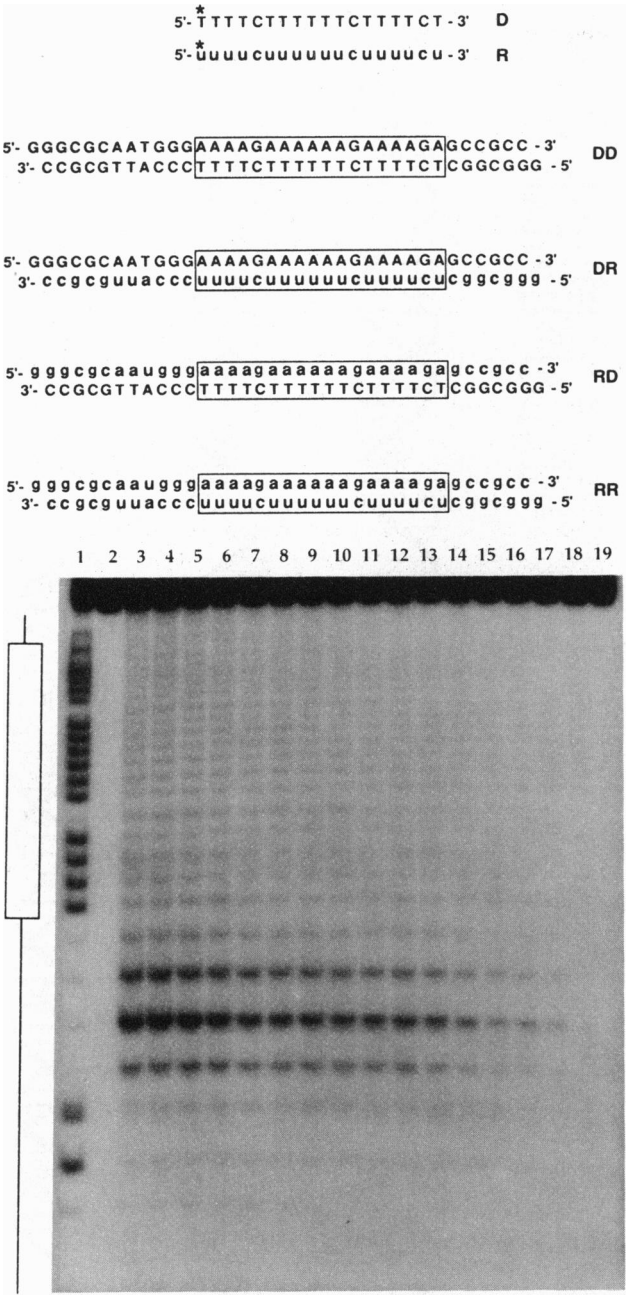

FIG. 2. (*Upper*) Sequences of oligodeoxyribonucleotide-EDTA (D), oligoribonucleotide-EDTA (R), and 35-bp duplexes DD, DR, RD, and RR. *T and *u are thymidine-EDTA and uridine-EDTA, respectively. The uppercase letters indicate deoxyribonucleotides (DNA) and the lowercase letters indicate ribonucleotides (RNA). The box indicates the 18-bp purine target sequence within the 35-bp duplex bound by D or R. (*Lower*) Autoradiogram of a 20% denaturing polyacrylamide gel revealing cleavage products generated by D·Fe in a quantitative affinity cleavage titration experiment performed on the duplex DD labeled at the 5′ end of the purine strand with ^{32}P. The white box drawn on the left of autoradiogram indicates the position of the 18-bp target site of the duplex DD. Lane 1, products of an A-specific cleavage reaction of DD (17). Lane 2, control showing intact ^{32}P-labeled duplex obtained after incubation under the conditions of cleavage reactions in the absence of D·Fe. Lanes 3–19, DNA affinity cleavage products produced by D·Fe at various concentrations: 10.0 μM (lane 3), 8.0 μM (lane 4), 4.0 μM (lane 5), 2.0 μM (lane 6), 1.0 μM (lane 7), 800 nM (lane 8), 400 nM (lane 9), 200 nM (lane 10), 100 nM (lane 11), 80 nM (lane 12), 40 nM (lane 13), 20 nM (lane 14), 10 nM (lane 15), 8 nM (lane 16), 4 nM (lane 17), 2 nM (lane 18), 1 nM (lane 19).

each of four 35-bp duplexes, DD, DR, RD, and RR, were measured by quantitative affinity cleavage titration (Fig. 2 *Upper*). For this, 5′-^{32}P-end-labeled duplex DNA or RNA and various concentrations of an oligonucleotide-EDTA·Fe D (1 nM–10 μM) or R (1 nM–40 μM) were allowed to equilibrate for 24 hr at 25°C in association buffer. Then, the cleavage reactions were initiated by addition of dithiothreitol and incubated for 8 hr at 25°C. The results of a representative experiment performed employing oligonucleotide-EDTA·Fe D targeting the duplex DD are shown in Fig. 2 *Lower*.

The amount of product fragments produced during an affinity cleavage experiment is proportional to the fractional occupancy of the duplex target site by the oligonucleotide-EDTA·Fe (11). By measuring the site-specific cleavage produced by bound oligonucleotide-EDTA·Fe as a function of total oligonucleotide concentration, an empirical titration binding isotherm can be constructed and the equilibrium association constant (K_T) can be determined (11). The amounts of radiolabeled DNA in the bands at the target cleavage site and at a reference site were measured from a photostimulable storage phosphor autoradiogram (18). The I_{site} data points were calculated according to Eq. 1 and the ($[O]_{tot}$, I_{site}) data points were fitted using a nonlinear least squares method, with K_T and I_{sat} as adjustable parameters (Eq. 2). The mean values of the association constants and the corresponding free energies for D and R binding with DD, DR, RD, and RR were each extracted from three such experiments (Table 1; see Fig. 4). The data points obtained were averaged and plotted with average best-fit titration binding isotherms (Fig. 3). Triple helix formation of D + RD and D + RR was not observed up to 40 μM concentration of D. Therefore, only upper limits on these two association constants can be estimated. The relative stabilities of the triple helical complexes are D + DD, R + DD, R + DR, D + DR > R + RD, R + RR >> D + RR, D + RD.

DISCUSSION

DNA vs. RNA. The differences in the energetics and structures of deoxyribonucleotide and ribonucleotide triple helices undoubtedly are related to the different chemical nature of DNA and RNA. For DNA there is a methyl group at the 5 position of the pyrimidine, thymine. For RNA, the methyl group is absent on uracil. On RNA there is a hydroxyl group at the 2′ position of the ribose sugar that is absent on DNA. One would expect the stacking of a methyl substituent (or lack of) on the adjacent pyrimidine in the bound third strand of the triple helix to be important. It is known from previous work that oligodeoxyribonucleotides containing 2′-deoxyuridine binds more weakly to duplex DNA by triple helix formation than oligodeoxyribonucleotides containing thymidine (19). Hence, we would expect the change of T to U to cause a decrease in affinity, likely due to loss of a stacking interaction. The presence or absence of the 2′-hydroxyl on the sugar that distinguishes RNA from DNA likely determines the conformational preference for A or B form with C-3′ endo and C-2′ endo sugar conformations, respectively. RNA duplexes prefer to adopt an A-form helical structure (20). Polymorphic DNA can adopt B or A form, and the B–A equilibrium is likely sequence composition dependent (20). Hence, the 2′-OH substituents will exert conformational preferences on the triple helical structure(s). X-ray diffraction and NMR studies of DNA, RNA, and hybrid duplexes and triplexes reveal that sugar conformation can differ in individual strands (21–26).

We found that six stable triplexes are formed, which is consistent with results of Roberts and Crothers (14). However, the order and the scale of free energy values differ (Fig. 4) (14). When the Watson–Crick duplex contains RNA in the purine strand, RNA, but not DNA, will bind with reasonable

Table 1. Equilibrium association constants (K_T, M^{-1}) measured by quantitative affinity cleavage titration at 25°C and pH 7.0

Hoogsteen strand*		Watson–Crick duplex	Hoogsteen strand†	
D	R	(Pu·Py)	D	R
3.9 (±0.7) × 10^6	3.3 (±1.3) × 10^6	DD	4.2 (±0.6) × 10^6	5.6 (±0.5) × 10^6
2.2 (±0.3) × 10^6	2.9 (±1.1) × 10^6	DR	2.0 (±0.6) × 10^6	3.2 (±1.0) × 10^6
<10^4‡	6.1 (±1.0) × 10^5	RD	<10^4‡	7.1 (±1.2) × 10^5
<10^4‡	5.4 (±1.0) × 10^5	RR	<10^4‡	9.2 (±2.2) × 10^5

Pu, purine; Py, pyrimidine.

*Hoogsteen strand binding to the duplex labeled at the 5′ end of the purine Watson–Crick strand with ^{32}P.

†Hoogsteen strand binding to the duplex labeled at the 5′ end of the pyrimidine Watson–Crick strand with ^{32}P.

‡These estimated values are based on no detectable cleavage of the duplexes RD and RR by D·Fe up to 40 μM.

affinity. The six binding affinities observed for this 18-nt sequence studied at pH 7.0 and 25°C were clustered and ranged from 3.9 × 10^6 to 5.4 × 10^5 M^{-1} (Table 1). The corresponding free energies are −9.1 and −7.9 kcal·mol^{-1} and differ only by 1.2 kcal·mol^{-1}. We estimate that the two possible triplexes that do not form (D + RR and D + RD) must be ≥3.6 kcal·mol^{-1} less stable than the R + DD complex. In contrast, Roberts and Crothers (14) measured more than 10 kcal·mol^{-1} variation (corresponding to a 10^8-fold difference in binding affinity) separating the most stable and least stable of the six complexes by van't Hoff analysis of melting curves of six triplexes. We cannot rule out at this time that the differences arise from the different methods of analyses. However, the binding affinities of R or D third strands to double helical nucleic acids are known to be length, sequence composition, pH, and salt dependent (11–13). Therefore, this difference more likely arises from experimental differences in the two studies: length (18 bp vs. 12 bp), base composition (15 A·T, 3 G·C vs. 4 A·T, 8 G·C), pH (7.0 vs. 5.5), cation concentrations (1 mM vs. 0 mM spermine), and stability of the target duplex (35-bp duplex vs. 12-bp hairpin duplex with 4-nt loop).

Recognition of Double Helical DNA. In this study, the free energies of binding of D and R to double helical DNA are comparable (−9 kcal·mol^{-1}). Oligonucleotides containing D or R will be useful for targeting double-stranded DNA by triple helix formation. This suggests that *in vivo* targeting of double helical DNA with RNA may be possible. Since the uracil does not contribute as much as thymine to binding, we conclude that the 2′-OH on RNA is compensating in a

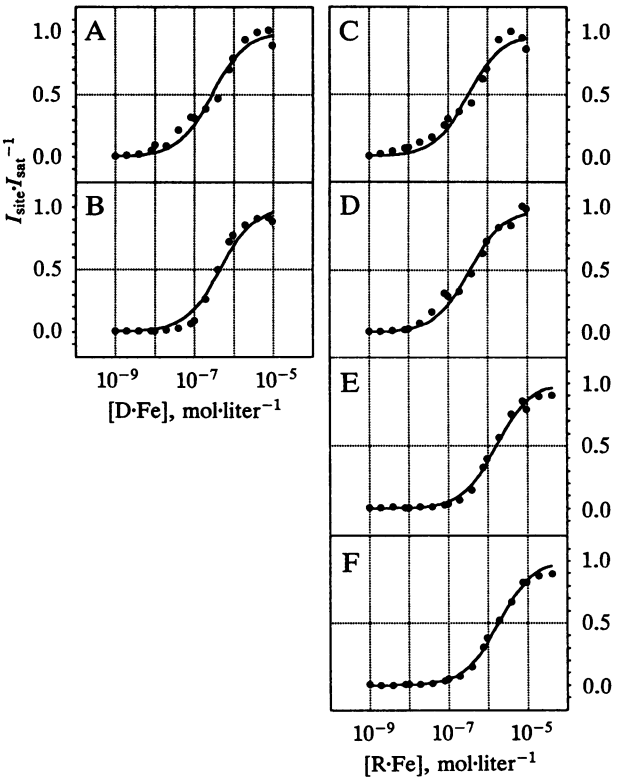

FIG. 3. Binding isotherms derived from quantitative affinity cleavage titration for six triple helical complexes formed. The I_{site}/I_{sat}^{-1} data represent the average of normalized site-specific cleavage signal intensities from three experiments and binding curves were plotted using the mean values of K_T. (*A–F*) Binding curves obtained for the duplexes DD to RR labeled at the 5′ end of the purine strand with ^{32}P: D + DD (*A*), D + DR (*B*), R + DD (*C*), R + DR (*D*), R + RD (*E*), R + RR (*F*).

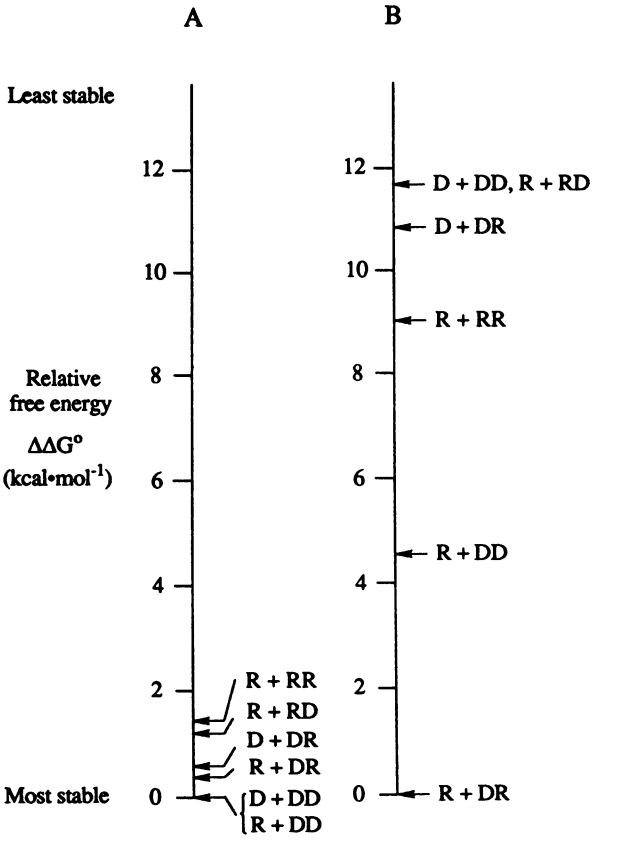

FIG. 4. Comparison of relative stabilities [ΔΔG^o, kcal·mol^{-1} (1 kcal = 4.18 kJ)] for six triple helical complexes taken from Table 1 (*A*) with values for recent report by Roberts and Crothers (*B*; ref. 14).

positive way. In the case studied here R and D bind DD with similar affinities, whereas Roberts and Crothers (14) found R formed a more stable complex, R + DD > D + DD (Fig. 4).

Recognition of Double Helical DNA·RNA and RNA. The free energies of binding to RNA·DNA hybrids depend on whether the purine tract is RNA or DNA. When the double helix is purineD·pyrimidineR, then R and D bind with comparable affinities ($\approx 2 \times 10^6$ M^{-1}). However, if the hybrid is purineR·pyrimidineD, then R, but not D, binds with similar affinity ($\approx 6 \times 10^5$ M^{-1}). Our results differ with those of Roberts and Crothers (14), in which R + DR was significantly more stable than D + DR. Finally, R, but not D, will bind double helical RNA. The free energy of R binding double helical RNA (R + RR) is $\approx$1.0 kcal·mol^{-1} less favorable than R binding double helical DNA (R + DD) or the DR hybrid (R + DR) but similar to the RD hybrid (R + RD).

Implications. It would appear that there could be several conformational families of triple helices related to the sugar identities of the three strands. Elucidation of the structures of these complexes must await direct studies such as x-ray or high-resolution two-dimensional NMR analysis. With regard to the sequence-specific recognition of double helical nucleic acids by triple helix formation, a simple guiding principle is the following: If the double helix contains DNA in the purine strand, RNA or DNA will bind; if the double helix contains RNA in the purine strand, only RNA will bind.

We are grateful to the National Institutes of Health for grant support. We thank K. Breslauer and D. Crothers for helpful discussions.

1. Moser, H. E. & Dervan, P. B. (1987) *Science* **238**, 645–650.
2. LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.-L., Thuong, N. T., Lhomme, J. & Helene, C. (1987) *Nucleic Acids Res.* **15**, 7749–7760.
3. Maher, L. J., III, Wold, B. & Dervan, P. B. (1989) *Science* **245**, 725–730.
4. Strobel, S. A. & Dervan, P. B. (1991) *Nature (London)* **350**, 172–174.
5. Strobel, S. A., Doucette-Stamm, L. A., Riba, L., Housman, D. E. & Dervan, P. B. (1991) *Science* **254**, 1639–1642.
6. Felsenfeld, G., Davies, D. R. & Rich, A. (1957) *J. Am. Chem. Soc.* **79**, 2023–2024.
7. Rajagopal, P. & Feigon, J. (1989) *Biochemistry* **28**, 7859–7870.
8. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241**, 456–459.
9. Beal, P. A. & Dervan, P. B. (1991) *Science* **251**, 1360–1363.
10. Radhakrishnan, I., de los Santos, C. & Patel, D. J. (1991) *J. Mol. Biol.* **221**, 1403–1418.
11. Singleton, S. F. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114**, 6957–6965.
12. Singleton, S. F. & Dervan, P. B. (1992) *Biochemistry* **31**, 10995–11003.
13. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, T. & Helene, C. (1992) *Biochemistry* **31**, 9269–9278.
14. Roberts, R. W. & Crothers, D. M. (1992) *Science* **258**, 1463–1466.
15. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15**, 8783–8798.
16. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Plainview, NY), 2nd Ed.
17. Iverson, B. L. & Dervan, P. B. (1987) *Nucleic Acids Res.* **15**, 7823–7830.
18. Johnston, R. F., Pickett, S. C. & Barker, D. L. (1990) *Electrophoresis* **11**, 355–360.
19. Povsic, T. J. & Dervan, P. B. (1989) *J. Am. Chem. Soc.* **111**, 3059–3061.
20. Saenger, W. (1984) *Principles of Nucleic Acids Structure* (Springer, New York).
21. Wang, A. H.-J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, S. A. A. & Rich, A. (1982) *Nature (London)* **299**, 601–604.
22. Arnott, S., Bond, P. J., Selsing, E. & Smith, P. J. C. (1976) *Nucleic Acids Res.* **3**, 2459–2470.
23. Chou, S.-H., Flynn, P. & Reid, B. (1989) *Biochemistry* **28**, 2435–2443.
24. Arnott, S., Chandrasekaran, R., Millane, R. P. & Park, H.-S. (1986) *J. Mol. Biol.* **188**, 631–640.
25. Macaya, R. F., Schultze, P. & Feigon, J. (1992) *J. Am. Chem. Soc.* **114**, 781–783.
26. Macaya, R. F., Wang, E., Schultze, P. & Feigon, J. (1992) *J. Mol. Biol.* **225**, 755–773.