Preferential Interaction of the mRNA Proofreading Factor TFIIS Zinc Ribbon with rU·dA Base Pairs Correlates with Its Function†

HoSup Yoon, Albert S. Sitikov, ChoonJu Jeon, and Kan Agarwal*

Departments of Biochemistry & Molecular Biology and Chemistry, The University of Chicago, Chicago, Illinois 60637

Received April 23, 1998; Revised Manuscript Received July 6, 1998

ABSTRACT: The transcriptional factor TFIIS helps overcome elongation barriers and enhances proofreading by RNA polymerase II. These TFIIS functions may be modulated by the TFIIS zinc ribbon domain through interactions with nucleic acids in the elongation complex. Within this zinc ribbon domain, the dipeptide sequences Asp261-Glu262 and Arg276-Trp277 have been shown to be critical for its function by mutant analysis. The sequence Asp261-Glu262 has been suggested to participate in metal binding within the RNA polymerase II active site. We now show that the sequence Arg276-Trp277 interacts with nucleic acids through a combination of electrostatic and stacking interactions. The interaction of the indole side chain of the tryptophan residue with nucleic acid bases is demonstrated by a characteristic and reversible decrease in the zinc ribbon fluorescence intensity as a function of oligonucleotide concentration. These interactions are salt sensitive (maximum interaction at 200 mM and no interaction at 500 mM NaCl), suggesting that the tryptophan stacking with nucleic acid base accompanies electrostatic contacts. The oligonucleotide−zinc ribbon interactions exhibit small but significant base preferences, as shown by the dependence of K_{eq} on base composition, with decreasing K_{eq} in the order U > T > A > C ≫ G. Within the variety of homopolymeric single- and double-stranded deoxy- and ribooligonucleotides, the oligonucleotide rU_{12-18}·dA_{20} exhibited a 2−6-fold binding preference relative to other oligonucleotides. This preferential binding of the zinc ribbon to sequences composed of rU·dA base pairs, which are generally associated with elongation blocks, may help in overcoming elongation barriers. Since the mRNA proofreading and enhancement of elongation involve cleavage of ribonucleotide of the mismatched pair and the weakly paired rU·dA nucleotides, but not the stably paired rC·dG nucleotides, we propose that the Arg276-Trp277 sequence in the TFIIS zinc ribbon may serve as a scanner connected to the transcript cleavage apparatus for weakly paired or mismatched nucleotides by employing indole ring stacking with the bases as a criterion of determining their subsequent removal. The striking similarity in preference for mismatched and weakly paired nucleotides for binding and for excision suggests a functional relationship between binding and cleavage reactions.

Messenger RNA synthesis by RNA polymerase II (pol II)[1] is regulated at all phases of transcription: RNA chain initiation, elongation, and termination. In the elongation process, pol II can encounter barriers leading to trapped RNA polymerase−RNA−DNA ternary complexes which are able to neither elongate nor dissociate and consequently prevent synthesis of full-length mRNA (*1*, *2*). A number of factors have been identified which modify pol II read through at such impediments. These include TFIIF, P-TEFb, Elongin SIII, ELL, and TFIIS (*2*). Of these factors, TFIIS is the best characterized elongation factor. It stimulates elongation by overriding barriers (*3*−*5*) and exhibits a novel ability to stimulate pol II 3′ → 5′ ribonuclease activity (*6*−*8*).

Two major structural domains specify TFIIS function: the C-terminal domain (residues 175−280), including the zinc ribbon motif, which is responsible for binding to nucleic acids, and the middle domain of similar size, which binds to pol II (*4*). Nucleic acid binding is not exhibited by the intact TFIIS protein but is unmasked by deletion of the pol II binding region (N-terminal residues 1−174 or central residues 165−174). In the intact protein, several point mutations in the zinc ribbon have been identified that reduce stimulatory ribonuclease activity (*8*). The fact that the isolated nucleic acid binding zinc ribbon does not bind to pol II demonstrates that protein sequences individually required for pol II binding and nucleic acids are jointly required for activity (*4*). The zinc ribbon is folded in the native protein to provide a cryptic nucleic acid binding surface which may be exposed on binding of TFIIS to an elongation complex (*4*). A similarly masked architecture of the nucleic acid binding domain has been found for the *Escherichia coli* RNA polymerase σ^{70} subunit (*9*). In each case, assembly-dependent exposure of functional surfaces may provide an allosteric mechanism for restricting sites of protein−nucleic acid interaction to the initiation (σ^{70}) or elongation (TFIIS) complex. Such regulation would circumvent the problem of nonspecific nucleic acid binding.

† This work was supported by U.S. Public Health Service Grant DK-21901.

* To whom correspondence should be addressed.

[1] Abbreviations: pol II, RNA polymerase II; Δ, deletion; ▽, insertion; dsDNA, double-stranded DNA; ssDNA or -RNA, single-stranded DNA or RNA; eZR, extended zinc ribbon (residues 175−280); ZR, zinc ribbon (residues 230−280; *4*).

S0006-2960(98)00924-6 CCC: $15.00
Published on Web 08/08/1998

TFIIS appears to enhance pol II-catalyzed RNA cleavage activity by modifying its active site (*10*). A previous finding that the sequence Asp261-Glu262 in the disordered loop of the zinc ribbon is required for enhancement of pol II ribonuclease activity suggested that a mechanism similar to that proposed for 3′ → 5′ exonuclease activity for the Klenow fragment of DNA polymerase I may operate here as well (*8*, *11*). Indeed, the recent finding that three acidic residues in the pol II active site are critical for its function lends support to this proposal (*12*). Moreover, several residues in the third β-strand of the zinc ribbon have also been identified as critical residues for TFIIS function by systematic site-directed mutagenesis. These studies revealed that the sequence Arg276-Trp277 is required for TFIIS function (*13*, *14*).

Here we explore the mechanism by which the dipeptide Arg276-Trp277 in TFIIS might enhance pol II ribonuclease activity. On the basis of the nature of the basic−aromatic nature of the dipeptide, we reasoned that these residues might participate in protein−nucleic acid interactions through a combination of electrostatic and stacking interactions. To test this proposal, we have used fluorescence spectroscopy to determine binding of these residues with single- and double-stranded oligodeoxynucleotides, single- and double-stranded oligoribonucleotides, and RNA−DNA hybrids. The data show that the dipeptide Arg276-Trp277 interacts with the single-stranded and double-stranded deoxy- and ribo-homopolymers and RNA−DNA hybrids. Within these homopolymers, a significant preference for the $rU_n{\cdot}dA_n$ homopolymer is exhibited. These results are discussed with respect to the specificity of interaction and their relationship to both suppression of pause sites and enhancement of proofreading activity.

EXPERIMENTAL PROCEDURES

Homopolymeric Oligonucleotides. Oligonucleotides $(dA)_{20}$, $(dC)_{10}$, $(dG)_{10}$, $(rC)_9$, $(dT)_{19-24}$, and $(rA)_{12-18}$ were purchased from P-L Biochemicals. The $(rU)_{12-18}$ was prepared by partial hydrolysis of poly(U) (obtained from Pharmacia) by 0.5 M NaOH, followed by size purification by polyacrylamide/7 M urea gel electrophoresis. Homopolymers $(dT)_{19-24}$, $(rA)_{12-18}$, and $(rU)_{12-18}$ contained greater than 80% dT_{19}, rA_{12}, and rU_{12}, respectively, as estimated by gel electrophoresis. Oligonucleotides $(dU)_{15}$, $(rU)_{15}$, $(dT)_{15}$, and $(dA)_{15}$ were chemically synthesized according to the manufacturer's protocol. The deprotected oligonucleotides were purified by gel electrophoresis. The oligonucleotide concentration was estimated spectrophotometrically using extinction coefficients at their wavelength maxima.

Proteins. The His_6-tagged extended zinc ribbon (residues 175−280) and zinc ribbon (residues 230−280) fusion proteins were expressed in *E. coli* and purified by Ni^{2+}-affinity chromatography as previously described (*14*, *15*). The proteins were >95% homogeneous as judged by SDS−PAGE. Protein concentrations were determined from the absorbance at 280 nm (extinction coefficient $\epsilon_{280} = 7.19 \times 10^3\ M^{-1}\ cm^{-1}$ for ZR and $\epsilon_{280} = 7.27 \times 10^3\ M^{-1}\ cm^{-1}$ for eZR).

Fluorescence Measurements. The binding buffer used for fluorescence measurements consisted of 10 mM sodium phosphate, 100 mM NaCl, 0.1 mM DTT, and 0.02 mM EDTA, adjusted to pH 7.4. Fluorescence measurements were carried out at 25 °C, on a Photon Technology International LPS-220 Spectrofluorometer using a 50 μL (3 mm × 3 mm) microcuvette. The concentration of the protein was calculated on the basis of its monomer form in solution (*15*). The excitation wavelength was 278 nm, and the fluorescence intensity was measured over the specified range. To reduce the absorbance by the oligonucleotides and minimize inner filter corrections, a mixture of mononucleotides with the same base composition and concentration was used along with oligonucleotides (*16*). The concentrations of total protein and oligonucleotides at each point in titration are determined after accounting for dilution.

After a protein solution had been titrated with an oligonucleotide to a maximum fluorescence quenching, a salt back-titration was performed. Small aliquots of the high-salt buffer were added to this preformed protein−oligonucleotide complex until approximately 90% of the original fluorescence was achieved.

RESULTS

Previously, we showed binding of the extended zinc ribbon (residues 175−280, herein abbreviated as eZR) to dsDNA, ssDNA, RNA, and RNA−DNA oligonucleotides by employing gel mobility shift assays (*4*). While these and mutant data suggested that binding involved the zinc ribbon motif (residues 230−280, herein abbreviated as ZR) within the eZR, they did not directly demonstrate its involvement in binding. Direct binding of the ZR to the same oligonucleotides that were bound by the eZR could not be reproducibly demonstrated by this method (*17*). The lack of mobility shift exhibited by the ZR could have been due to the ZR (50 residues) being only half the size of the eZR (111 residues). This relatively small size peptide may have a lower ZR−oligonucleotide complex stability that cannot be detected by the gel mobility shift method. To demonstrate binding of ZR to nucleic acids, an alternative, more sensitive, and direct method is needed.

The preferential binding of the eZR to pyrimidine-rich sequences suggested involvement of a motif that participated in discrimination between pyrimidines and purines (*4*). On the basis of earlier model studies which demonstrated that the tripeptide KWK interacted preferentially with pyrimidine-rich sequences, we predicted that the highly conserved sequence Arg276-Trp277-Lys278 in TFIIS might participate in interactions with nucleic acids (*18*, *19*). Indeed, mutant analysis of this segment convincingly demonstrated a critical role for the sequence Arg276-Trp277 in TFIIS function (*13*, *14*). Interestingly, residues Arg276 and Trp277 can only be substituted by other basic and aromatic residues, respectively, suggesting the possible involvement of charge−charge attraction and aromatic side chain intercalation with nucleic acid bases. If the indole side chain of Trp277 stacks between bases, as has been suggested for the KWK model peptide, such interactions can be quantitatively monitored by measuring changes in tryptophan fluorescence emission (*19*). Therefore, to test this hypothesis, we determined the tryptophan fluorescence emission of eZR and ZR in the presence of various oligonucleotides.

Fluorescence Emission Spectra of TFIIS, the Extended Zinc Ribbon (eZR), and the Zinc Ribbon (ZR). The excitation

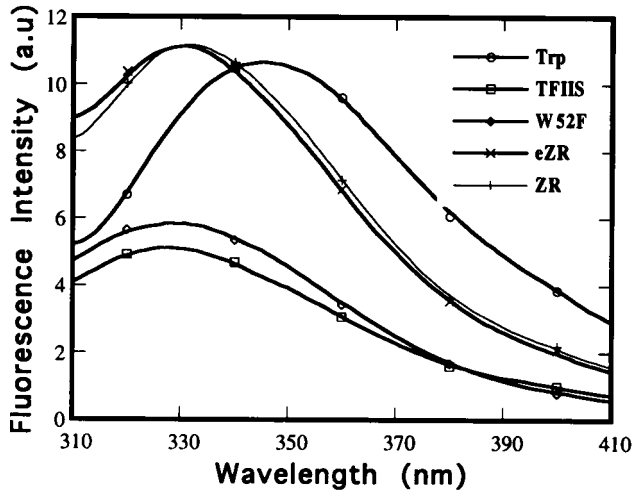

FIGURE 1: Fluorescence emission spectra of TFIIS, mutant W52F, the truncated extended zinc ribbon (eZR), and the zinc ribbon (ZR) relative to that of tryptophan. The excitation wavelength was 278 nm, and the emission spectra were recorded from 310 to 400 nm. The concentration of each of the proteins was 2 μM, and that of L-Trp was 1 μM.

wavelength selected for the spectra was 278 nm. At this wavelength, both tryptophan and tyrosine can exhibit fluorescence emission. However, the fluorescence emission of most native proteins is dominated by the tryptophan fluorescence, and tyrosine fluorescence is relatively small and frequently undetectable (*20*). The single tyrosine 252 in eZR and ZR does not significantly contribute to the tryptophan fluorescence emission and to the TFIIS function because this residue can be substituted by leucine without a change in the spectrum and TFIIS activity (*13*).

The TFIIS contains two tryptophan residues, W52 and W277. Residue W277 is required for activity, whereas the amino-terminal W52 is not (*14*). To determine the local environment of the tryptophan residues, fluorescence emission spectra of TFIIS, eZR, ZR, and the W52F mutant were recorded. The 20 nm blue shift relative to tryptophan exhibited by TFIIS indicates that both residues W52 and W277 lie in a similar hydrophobic environment (Figure 1). Similar analysis of the mutant W52F showed no change in fluorescence maxima and thus supported the above conclusion. Since W277 may be involved in nucleic acid binding, its environment was determined by analysis of the fluorescence emission of eZR and ZR. In both proteins, residue W277 lies in a comparable hydrophobic environment as evidenced by the same 20 nm blue shift (Figure 1). These findings are further supported by the NMR studies with ZR which showed the indole side chain of the tryptophan is internally located (*15*). From these data, we conclude that the structure of the zinc ribbon in ZR and eZR is similar.

Quenching of eZR and ZR Fluorescence by Homopolymeric Oligonucleotides. To determine base-specific binding of eZR, we employed deoxy- and ribooligonucleotide homopolymers representing each of the nucleic acid bases as single- and double-stranded DNAs, single-stranded RNAs, and RNA−DNA hybrids averaging 15−20 nucleotides in length. If zinc ribbon W277, which is required for TFIIS function, as has been suggested by analysis of Arg276-Trp277-Lys278 region mutants, interacts with nucleic acid bases through intercalation, then tryptophan fluorescence should be quenched by the bound nucleic acid base. The relative level of quenching should depend on the avidity of the zinc ribbon tryptophan for a particular base and its accessibility in a given structural environment. To determine whether the zinc ribbon exhibits base-specific interactions, fluorescence emission spectra of eZR were measured for single-stranded homopolymers as a function of oligonucleotide concentration. The small homopolymers are ideal ligands for base specificity determinations because of their structural simplicity.

Analysis of single-stranded homopolymers $(dA)_{20}$, $(dG)_{10}$, $(dC)_{10}$, $(dT)_{19-24}$, $(dU)_{15}$, $(rC)_9$, and $(rA)_{12-18}$ and the eZR protein showed consistently an oligonucleotide concentration-dependent decrease in fluorescence intensity at all wavelengths. The homopolymer $(rG)_{10}$ posed solubility problems and therefore could not be analyzed. A few selected fluorescence emission spectra of eZR with homopolymers $(dA)_{20}$, $(dT)_{19-24}$, and $(rU)_{12-18}$ are presented in Figure 2A as examples. As shown in Figure 2A, the fluorescence quantum yield of 2.0 μM eZR is dependent on the oligonucleotide concentration, reaching a minimum of approximately 50% of the unquenched value (i.e., 50% quenching). At higher oligonucleotide concentrations (2.0−50 μM), for most oligonucleotides, the quenching reached a maximum of 80%, but with oligonucleotides $(dG)_{10}$ and $(rC)_9$, it reached only 60% (data not shown). Similar levels of quenching were observed with ZR under comparable concentrations of protein and oligonucleotide (data not shown).

Double-stranded homopolymers $(dA)_{20}\cdot(dT)_{19-24}$, $(dC)_{10}\cdot(dG)_{10}$, $(dG)_{10}\cdot(rC)_9$, $(dT)_{19-24}\cdot(rA)_{19-24}$, $(dU)_{15}\cdot(dA)_{15}$, and $(dA)_{20}\cdot(rU)_{12-18}$ also showed a marked decrease in eZR fluorescence intensity at 335 nm with increasing oligonucleotide concentration. Selected fluorescence spectra of eZR with $(dA)_{20}\cdot(dT)_{19-24}$, $(dA)_{20}\cdot(rU)_{12-18}$, and $(dC)_{10}\cdot(dG)_{10}$, as examples, are presented in Figure 2B. Similar fluorescence quenching of ZR by oligonucleotides $(dA)_{15}\cdot(dT)_{15}$, $(dA)_{15}\cdot(rU)_{15}$, and $(dA)_{15}\cdot(dU)_{15}$ was observed, as shown in Figure 2C. We selected A, T, and U base-containing oligonucleotides for ZR studies because these homopolymers showed an increased level of ZR fluorescence quenching relative to C- and G-containing oligonucleotides (data not shown).

The decrease in protein fluorescence intensity at 335 nm (emission maxima) as a function of oligonucleotide concentration has been attributed to the $\pi-\pi$ stacking interactions between the indole side chain of tryptophan residue and the nucleic acid base (*16*, *21*, *22*). Therefore, we suggest that binding of eZR and ZR to single- and double-stranded oligonucleotides involves interaction of the indole side chain of Trp277 with the nucleic acid base. Since the single tryptophan residue at position 277 in eZR and ZR is located in the third β-strand of the zinc ribbon (*15*), we conclude that binding of eZR and ZR to oligonucleotides occurs primarily through the zinc ribbon motif.

Reversibility of eZR and the $(dA)_{20}\cdot(rU)_{12-18}$ Complex. Since most nucleic acid−protein interactions have a significant electrostatic component, a salt back-titration is a useful criterion for measuring equilibrium binding affinity and specificity at a constant temperature and pH. Intrinsic fluorescence measurements of eZR over a pH range of 4.0−9.0 revealed that, between pH 6.5 and 8.0, tryptophan fluorescence neither significantly shifted nor was quenched. Therefore, a near-neutral pH of 7.4 was chosen for assays

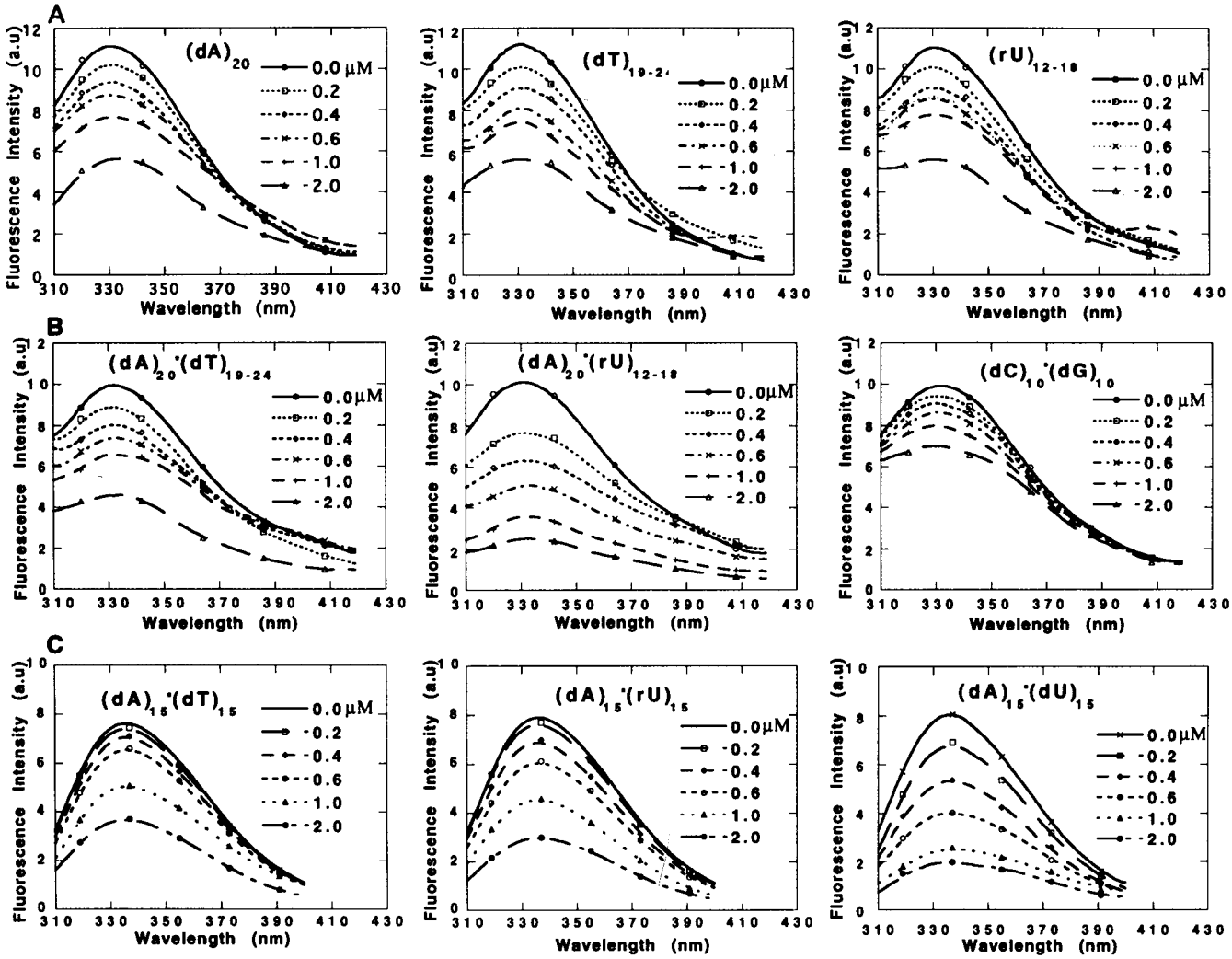

FIGURE 2: Fluorescence emission spectra of eZR and ZR (2.0 μM monomer) titrated with single- and double-stranded homopolymeric oligonucleotides. The oligonucleotide concentration is indicated in each panel. The excitation wavelength was 278 nm. The emission maximum was at 335 nm: (A) titration of eZR with single-stranded homopolymers, (B) titration of eZR with double-stranded homopolymers, and (C) titration of ZR with double-stranded homopolymers.

(*14*). Oligonucleotide $(dA)_{20}\cdot(rU)_{12-18}$ was chosen over the others because it showed a relatively higher level of fluorescence quenching at comparable concentrations (Figure 2B). Since 2 μM eZR bound with sufficient affinity to 0.4 μM $(dA)_{20}\cdot(rU)_{12-18}$ at 100 mM NaCl (Figure 2B), a complex formed at this concentration was titrated with increasing NaCl concentrations. Interestingly, the optimal binding of eZR required 200 mM NaCl (Figure 3). A similar salt optimum was estimated by the gel mobility shift assay (*4*). The preformed complex gradually dissociates with a further increase in salt concentration; at 500 mM NaCl, the quenched fluorescence is approximately 90% regenerated as intrinsic fluorescence (Figure 3). The salt-dependent dissociation of the eZR−oligonucleotide complex suggests that interaction of the tryptophan side chain may involve electrostatic interactions in addition to stacking interactions.

eZR and ZR Proteins Exhibit Nucleic Acid Base-Specific Binding. Titration of 2.0 μM eZR and ZR with different oligonucleotides exhibits varying levels of fluorescence quenching at a given oligonucleotide concentration. For instance, 2.0 μM $(dA)_{20}\cdot(dT)_{19-24}$ resulted in 50% quenching (Figure 2C), whereas 2.0 μM $(dC)_{10}\cdot(dG)_{10}$ caused 30% quenching (Figure 2B). To compare the fluorescence

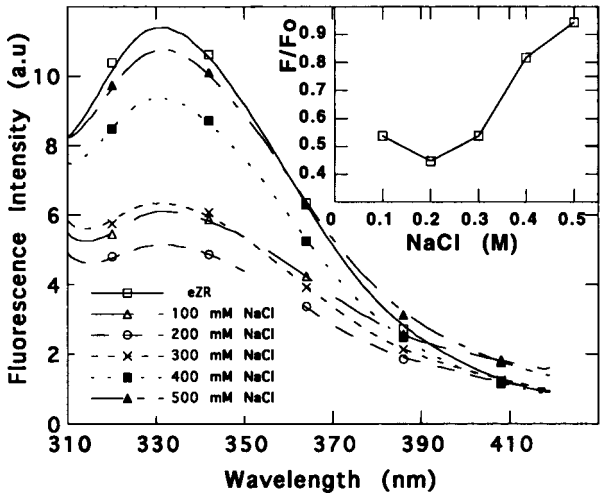

FIGURE 3: Fluorescence emission spectra of the eZR·$(dA)_{20}\cdot(rU)_{12-18}$ complex titrated with NaCl. The complex of 2.0 μM eZR with 0.4 μM $(dA)_{20}\cdot(rU)_{12-18}$ in 100 mM NaCl was titrated with increasing NaCl concentrations as indicated. The optimal concentration of NaCl for maximal fluorescence quenching is 200 mM (inset). Approximately 90% of the quenched fluorescence is regenerated as intrinsic fluorescence at 500 mM NaCl.

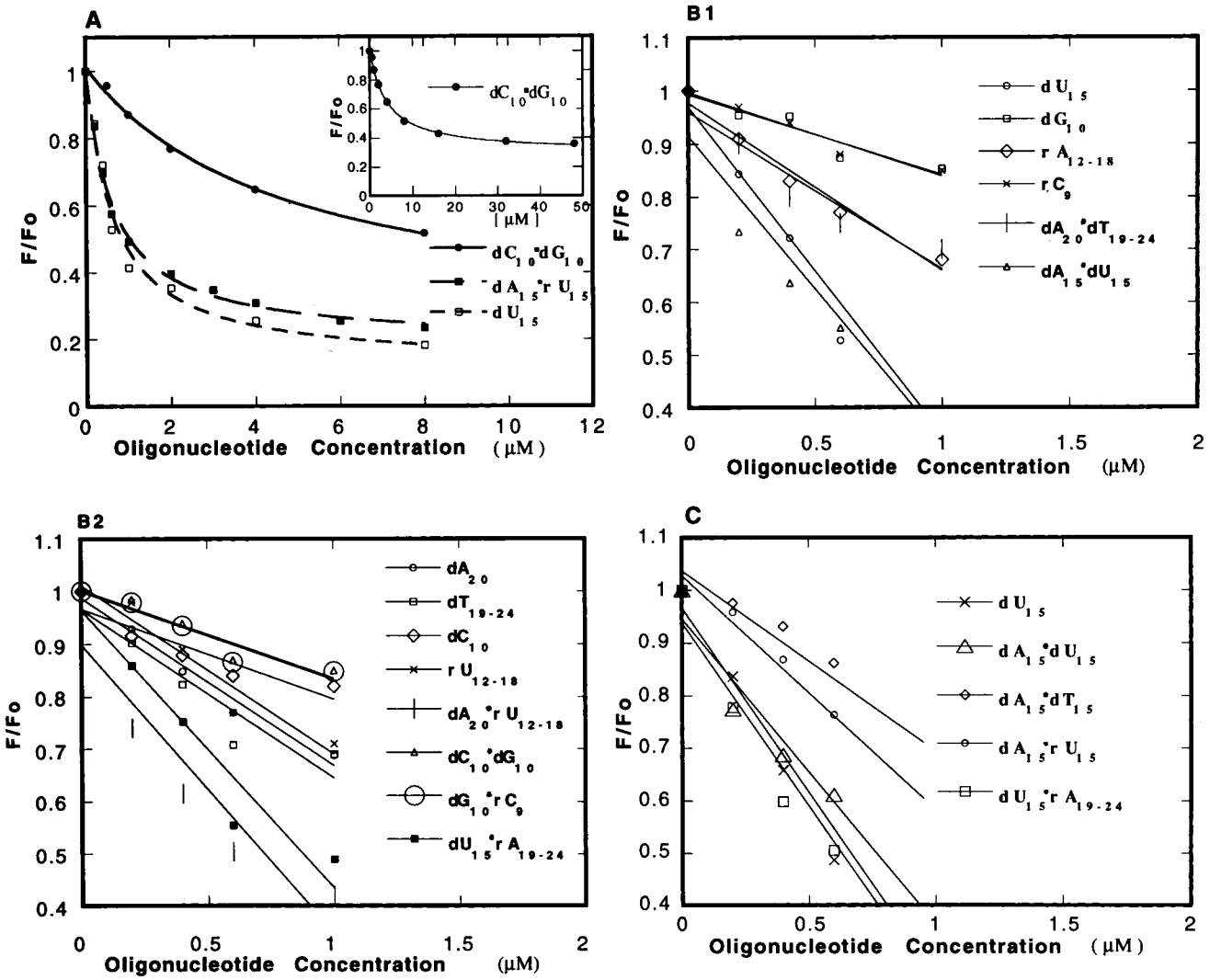

FIGURE 4: Binding isotherms of eZR and ZR with single- and double-stranded homopolymeric oligonucleotides. The data of fluorescence maxima (335 nm) from Figure 2 and from ref *14* were taken, and the values of F/F_o were plotted against the oligonucleotide concentration. F and F_o are fluorescence emission intensity maxima in the presence and absence of oligonucleotide, respectively. The inset shows the oligonucleotides that were examined. (A) Binding isotherms of eZR (2 μM) with low and higher oligonucleotide concentrations. (B1 and B2) Binding isotherms of eZR (2 μM) with low oligonucleotide concentrations. (C) Binding isotherms of ZR (2 μM) with low oligonucleotide concentrations.

quenching of eZR and ZR by oligonucleotides comprised of different bases, a plot of F/F_o (a fraction of the fluorescence remaining after the protein bound to the nucleic acid) versus an oligonucleotide concentration was constructed. The data presented in panel A of Figure 4, for selected oligonucleotides, demonstrate that the fluorescence quenching is linearly related to the fraction of the protein bound to the nucleic acids at low concentrations of oligonucleotides. Therefore, subsequent plots show binding to eZR (panels B1 and B2 of Figure 4) and ZR (panel C of Figure 4) up to an oligonucleotide concentration of 1.0 μM. From such a plot, a pattern of fluorescence decrease specific for a homopolymeric base begins to emerge (panels B1 and B2 of Figure 4). Among the single-stranded homopolymers, $(dT)_{19-24}$ and $(dU)_{15}$ showed the highest level of fluorescence quenching while $(rC)_9$ and $(dG)_{10}$ showed the least amount. The most preferred homopolymers were composed of dU, dT, dA, and rA, whereas the least preferred were composed of rU, dC, rC, and dG. While the influence of ribose versus deoxyribose is very small with A-based homopolymers, the significant difference between dU- and rU-based and between dC- and rC-based homopolymers suggests that eZR prefers deoxypyrimidines over ribopyrimidines. Within the pyrimidines, U is preferred over T or C.

This preference for U-based single-stranded homopolymers with the exception of rU-based oligonucleotide is extended to U-based double-stranded homopolymers. The eZR shows a preference for double-stranded homopolymers dA·rU-, rA·dU-, and dA·dU-based nucleotides over the rC·dG-based nucleotides (panels B1 and B2 of Figure 4). A similar preference is exhibited by the ZR (panel C of Figure 4), indicating that both eZR and ZR may interact through the common zinc ribbon Trp277 residue.

For quantitative analysis, the equilibrium constant K_{eq} was calculated by using a nonlinear least-squares fit (NLLSF) applied to plots of $F_o - F/F_o$ versus molar concentration of the oligonucleotide. Examination of the K_{eq} of eZR binding to single-stranded homopolymers (Table 1), which range from 0.64 μM for dU to 5.7 μM for dG, reveals a pattern of base specificity with the following hierarchy: dU > dT > dA ~ rA > rU ~ dC ≫ rC > dG. While we have not examined in detail the ZR−single-stranded homopolymer

Table 1: Equilibrium Constants (K_{eq}) for Peptide−Oligonucleotide Binding

oligonucleotide	eZR K_{eq}[a] (μM) (extended Zn ribbon)	ZR K_{eq}[a] (μM) (Zn ribbon)
$(dA)_{20}$	1.83 (±0.16)	
$(dC)_{10}$	2.67 (±0.35)	
$(dG)_{10}$	5.70 (±0.81)	
$(dT)_{19-24}$	1.46 (±0.10)	
$(dU)_{15}$	0.64 (±0.15)	0.63 (±0.14)
$(rA)_{12-18}$	1.75 (±0.13)	
$(rC)_9$	4.59 (±0.50)	
$(rU)_{12-18}$	2.37 (±0.40)	
$(dA)_{15}\cdot(rU)_{15}$		0.63 (±0.10)
$(dA)_{20}\cdot(rU)_{12-18}$	0.65 (±0.04)	
$(dG)_{10}\cdot(rC)_9$	3.49 (±0.68)	
$(dT)_{19-24}\cdot(rA)_{12-18}$	1.94 (±0.38)	
$(dU)_{15}\cdot(rA)_{19-24}$	0.68 (±0.15)	0.55 (±0.10)
$(dA)_{15}\cdot(dU)_{15}$	0.61 (±0.13)	0.55 (±0.07)
$(dA)_{15}\cdot(dT)_{15}$		1.86 (±0.16)
$(dA)_{20}\cdot(dT)_{19-24}$	1.88 (±0.39)	
$(dC)_{10}\cdot(dG)_{10}$	3.73 (±0.76)	

[a] Values of K_{eq} were obtained by using a nonlinear least-squares fit (NLLSF) program applied to plots of $(F_0 - F)/F_0$ vs oligonucleotide concentration. F and F_0 are fluorescence intensity maxima in the presence and absence of oligonucleotide at 335 nm, respectively. The values of F and F_0 are taken from Figure 2 and ref *14*.

interactions, the virtually identical K_{eq} of ZR (0.63 μM) and eZR (0.64 μM) for $(dU)_{15}$ suggests that the order of base and sugar specificity would be similar for these proteins.

Base specificity is also exhibited by the eZR in binding to double-stranded homopolymers. Specificity for dA·dU, dA·rU, and dA·T base pairs over C·G base pairs can be seen in Figures 2B and 4. The dA·dU homopolymer shows an approximately 6-fold higher affinity relative to C·G double-stranded homopolymers (compare K_{eq} for 0.61 μM for dA·dU to 3.73 μM for dC·dG; Figure 4 and Table 1). Interestingly, dA·dU, dA·rU, and rA·dU homopolymers have a similar affinity for eZR (K_{eq} ranging from 0.61 to 0.68 μM). However, this affinity is 3-fold higher than those of the dA·dT and rA·dT homopolymers (compare 0.61 μM for dA·dU with 1.88 μM for dA·dT and 1.94 μM for rA·dT). The virtually identical affinity of eZR for deoxyribose- and ribose-based A or U nucleotide suggests that eZR does not discriminate between deoxy- and ribo-based double-stranded homopolymers. At the single-stranded homopolymeric level however, discrimination between ribo- and deoxynucleotide is significant [compare K_{eq} for dU (0.64 μM) with that of rU (2.37 μM)].

It is possible that the 2′-OH group in single-stranded homopolymers reduces stacking of the eZR tryptophan residue, whereas in the double-stranded homopolymers, this influence is minimized. To explain this, we speculate that the 2′-OH group could potentially form a hydrogen bond with the 2-keto of pyrimidines in a single-stranded homopolymer, generating a distinct structure from that of a deoxynucleotide homopolymer. The vastly different electrophoretic mobility of dU_{15} and rU_{15} suggests that these oligonucleotides form different structures (data not shown). In purines, such hydrogen bonding cannot occur because the 2′-OH group cannot hydrogen bond with a purine base. In single-stranded ribo-homopolymer, such a hydrogen bond may restrict pyrimidine base movement which may be needed for proper stacking of the tryptophan. In double-stranded homopolymers, such hydrogen bonding would not occur because of the restricted rotation of the Watson−Crick paired pyrimidine base. This discrimination at the single-stranded level may reflect upon the mechanism by which the zinc ribbon may function in the elongation complex, i.e., whether it acts on the single-stranded nucleotide or on the double-stranded nucleotide.

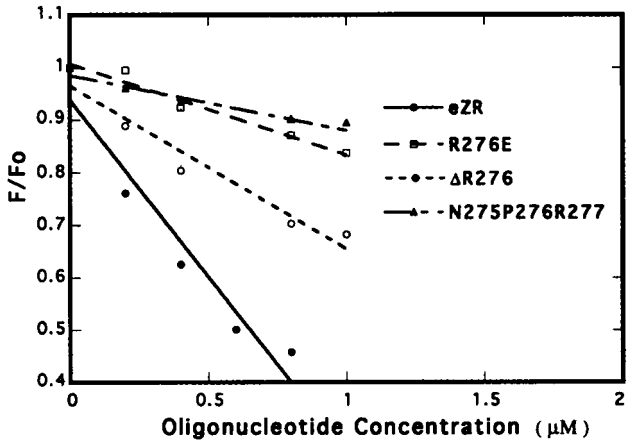

FIGURE 5: Binding isotherms of eZR mutants with $(dA)_{20}\cdot(rU)_{12-18}$. The fluorescence emission maxima from ref *14* were used to plot F/F_o against the oligonucleotide concentration. The inset shows the eZR mutants that were examined.

While we have not examined binding of ZR to as many homopolymers as we did with eZR, the binding affinity for selected homopolymers was similar to that of eZR. Both ZR and eZR exhibited specificity for A- and U-based double-stranded homopolymers (Figure 4 and Table 1). The strikingly similar binding affinities of ZR and eZR for these homopolymers further support the hypothesis that the zinc ribbon interacts with nucleic acids through intercalation of the tryptophan residue.

Residue R276 in R276W277 Is Critical for Activity. The importance of the basic residue at position 276, preceding W277, has previously been demonstrated by substitution, deletion, and insertion mutants (*13*). To evaluate the binding relationship to activity, three inactive mutants, R276E, ΔR276, and N275-P276-R277, were examined for fluorescence quenching by $(dA)_{20}\cdot(rU)_{12-18}$. In the mutants R276E and ΔR276, the hydrophobic environment was unchanged as evidenced by the fluorescence maximum (335 nm) that is identical to that of the wild type eZR (data not shown; *14*). However, a red shift to 345 nm indicated modification of the tryptophan environment in proline insertion mutant N275-P276-R277. These data suggest that substitution and deletion of R276 retains the hydrophobic environment of the indole side chain, presumably by maintaining the β-strand structure, whereas insertion of a proline residue exposes indole to a relatively hydrophilic environment, perhaps by loss of β-strand structure. The structural change caused by the P276 insertion is important for nucleic acid binding because this proline insertion mutant showed a 5−6-fold reduction in fluorescence quenching relative to eZR (Figure 5). A similar reduction in fluorescence quenching occurred for the mutant R276E, where the β-strand structure would be expected to be maintained but the acidic residue, preceding W277, might interfere in the binding to the nucleic acid backbone. This is supported by the mutant ΔR276 in which the preceding neutral residue, N275, weakens binding but does not prevent it. As a consequence, a 2−3-fold

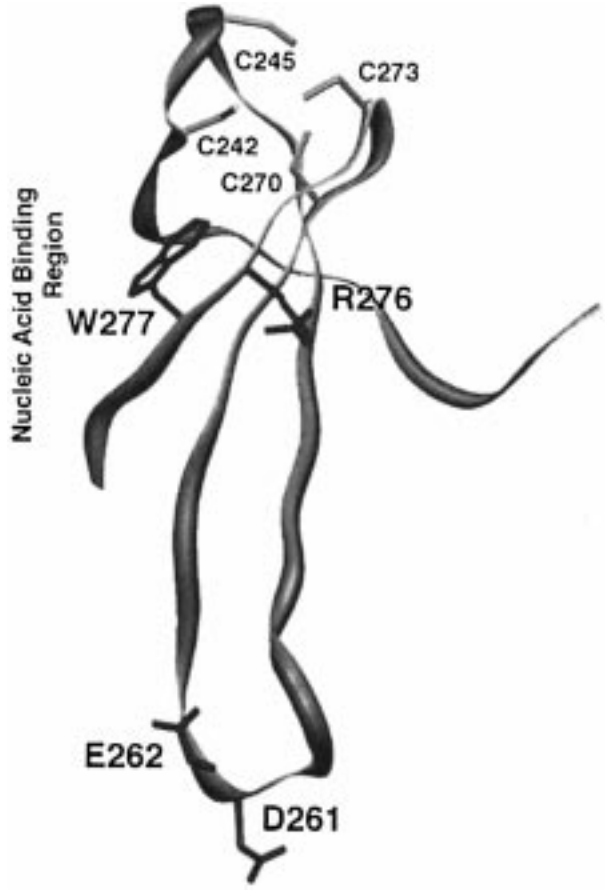

FIGURE 6: Structure of the TFIIS zinc ribbon (ZR) showing the explicit location of R276-W277 and D261-E262. The structure was determined by NMR (*15*). The four cysteines involved in zinc binding are orange.

reduction in fluorescence quenching occurs. From these results, we conclude that interaction of the tryptophan side chain with the nucleic acid is dependent on a preceding positively charged residue in a β-strand structure (Figure 6).

DISCUSSION

Since the discovery that TFIIS enhances pol II read through at elongation barriers (*3*−*5*), the major goal of our research has been to uncover the mechanism of TFIIS function. In our early analysis of TFIIS, we demonstrated that the extended zinc ribbon protein interacted with a variety of single- and double-stranded nucleic acids which led us to propose that the interaction of the zinc ribbon with the nucleic acid framework in a stalled elongation complex mitigated the transcriptional impediment (*4*). Subsequent studies identified two zinc ribbon dipeptides, Asp-Glu and Arg-Trp, which were required for both read through and transcript cleavage activities (*8*, *13*, *14*; Figure 6). The proposed role of the acidic residues of the dipeptide Asp-Glu was to participate in the phosphoryl transfer reaction via a two-metal ion mechanism (*8*, *11*). The dipeptide Arg-Trp was proposed to influence nucleic acid binding. However, the influences of substitution of these residues on the cleavage and nucleic acid binding (as measured by a gel retardation assay) in some cases were difficult to reconcile because of the qualitative nature of the nucleic acid binding data. For instance, the wild-type elongation stimulation activity of the R258K mutant could not be reconciled with the absence of a gel mobility shift with nucleic acids (*14*). Normally, the lack of a gel mobility shift by a mutant protein relative to the wild-type protein would be assumed to imply a deficiency in binding. However, it is possible that the weak intrinsic binding affinity ($\sim$2 μM) may have fallen below the range of measurement by the gel mobility shift assay, while retaining nucleic acid binding at a level sufficient for activity. This and other mutant studies demonstrated a need for a method for measuring nucleic acid binding which is more sensitive and reproducible than the gel mobility shift assay. Moreover, our intention was to determine if binding to nucleic acids could be quantitatively measured to reveal specificity in binding to a certain type(s) of nucleic acid base over others. By taking advantage of the single tryptophan residue (Trp277) of the critical dipeptide Arg-Trp in the zinc ribbon, we have been able to conduct a quantitative fluorimetric analysis of eZR and ZR binding with a variety of nucleic acids (Table 1).

The important role aromatic amino acids can play in binding of proteins to nucleic acids has been demonstrated (*18*, *19*, *23*, *24*). This investigation of the importance of the eZR sequence Arg276-Trp277 for interaction with nucleic acids of different structures and base composition sheds light on the nature of the interactions and reveals their direct relationship to TFIIS function. For example, the sensitivity of fluorescence quenching and TFIIS activity to increasing salt concentration is consistent with participation of arginine and tryptophan side chains in electrostatic and stacking interactions with nucleic acid phosphate and bases, respectively. Mutants R276E and ΔR276, which should exhibit reduced electrostatic interactions, showed decreased binding, as measured by a decrease in fluorescence quenching, and therefore are inactive. Thus, the interaction of ZR with polynucleotides is primarily through oppositely charged groups and stacking of an aromatic side chain with a base.

The binding of eZR with several single-stranded deoxy- and ribo-homopolynucleotides revealed base specificity in the following order of decreasing affinity: dU > dT > dA $\sim$ rA > rU > dC $\gg$ dG $\gg$ rC (panels B1 and B2 of Figure 4 and Table 1). A similar hierarchy has been observed for *E. coli* SSB and T4 gene 32 protein (both of which contain Trp) and mono-Trp oligolysines (*25*−*27*). For double-stranded oligonucleotides, a unique base specificity hierarchy of decreasing affinity is observed: dA·dU $\sim$ dA·rU $\sim$ rA·dU > dA·dT > rA·dT $\gg$ rC·dG $\gg$ dC·dG (panels B1 and B2 of Figure 4 and Table 1). Surprisingly, this double-stranded specificity is similar to the order of duplex stability studied by a variety of methods (*28*−*30*). This 4−5-fold binding preference for a weakly paired rU·dA oligonucleotide over the stably paired rC·dG oligonucleotide correlates remarkably well with TFIIS function. The TFIIS-stimulated pol II ribonuclease activity rapidly hydrolyzes weakly paired rU·dA nucleotides, whereas it is relatively less active for stably paired rC·dG nucleotides in a stalled elongation complex (*31*).

The transcriptional elongation is a dynamic process involving rapid association and dissociation of interacting molecules. Stimulation of pol II elongation by TFIIS could potentially involve enhancement by the zinc ribbon of the intrinsic rate of association and/or dissociation of pol II with nucleic acids in the elongation zone. The micromolar binding affinity of the zinc ribbon for nucleic acids could

represent a critical level at which association and dissociation of these molecules occur at a maximum efficiency. The lower binding affinity maintains high rates of association and dissociation of pol II-bound TFIIS with nucleic acids. Therefore, weak eZR binding with nucleic acids is mechanistically significant.

The function of TFIIS is to mitigate blocks to elongation and enhance proofreading (*2*, *31*). Both of these functions involve the weak rU·dA pairs (generally associated with the elongation blocks) and mismatched nucleotides. In the elongation zone, the TFIIS zinc ribbon may serve as a scanner connected to the transcript cleavage apparatus for weakly paired or mismatched nucleotides by employing indole ring stacking with the bases as a criterion of measurement. An increased ability to stack at a given site would identify site-specific duplex disruption because stacking of the indole ring at such sites would be energetically favored. At these sites, stacking may involve one of the two strands of the duplex because of the similar sizes of the indole ring and the purine base. Such stacking may induce a site-specific conformational change causing bending, kinking, or twisting of the duplex, leading to a favorable alignment of the phosphodiester moiety for metal-dependent hydrolysis. A similar mechanism has been proposed for the T4 UV endonuclease V and *E. coli* endonuclease III activities by employing the KWK peptide as a model. The specific binding of the KWK peptide at an apurinic site through indole ring stacking leads to formation of a Schiffs base between the lysine side chain amino group and the sugar aldehyde which promotes phosphodiester hydrolysis by β-elimination (*32*−*34*).

The zinc ribbon displayed specificity for A-, U-, and T-based double-stranded homopolymers with the following hierarchy: dA·dU ∼ dA·rU ∼ rA·dU > dA·dT (Table 1). Since nucleic acid architecture in the elongation complex involves elongating RNA hybridized to the transcribed DNA strand, the interaction of the zinc ribbon with dA·dU- and rA·dU-based homopolymers can be considered functionally irrelevant. The binding to dA·rU paired sequences is functionally significant because such sequences are generally associated with the elongation blocks that TFIIS can help overcome. Moreover, the TFIIS ribonuclease stimulatory activity exhibits a 40-fold preference for dA·rU base pairs over rC·dG base pairs (*31*), suggesting a functional relationship between nucleic acid binding by TFIIS and TFIIS-stimulated pol II cleavage activity.

Since rU·dA homopolymers preferentially exist as a triple helix, it is possible that a triple helix conformation exists in the stalled elongation zone (*35*). Could then the interaction of the zinc ribbon involve a triple helix rather than a duplex? While there is no direct evidence of the existence of a triple helix conformation in the elongation zone, some sequences prone to adopting a triple helix have been implicated in blocking transcription (*36*). If it is assumed that RNA polymerase II stalls because of the existence of a rU·dA triple helix, it is conceivable that stacking of the indole ring within the triple helix may induce kinking, resulting in disassembly of the triple helix into a double helix suitable for elongation resumption. A knowledge of the conformation of the nucleic acid framework in the stalled elongation zone is needed to test these ideas. Work on these lines is underway.

The participation of the ZR tryptophan in mitigation of an elongation block and in RNA cleavage was demonstrated by the site-directed mutagenesis studies (*13*). Substitution of tryptophan by phenylalanine generated a mutant protein with virtually identical activity. However, substitution by tyrosine resulted in reduction of read through activity (*13*). The reduced read through activity of the tyrosine mutant could result from the relatively altered stacking of the tyrosine with respect to the tryptophan in its interaction with the rU·dA pairs. This can be explained by a "bookmark" hypothesis which states that specificity for weakly stacked base pairs (such as rU·dA) decreases in the order tryptophan > phenylalanine > tyrosine > leucine (*18*). This marked functional variance exhibited by different aromatic side chains supports the contention that the TFIIS zinc ribbon can directly act on the nucleic acid framework and influence read through and proofreading activities.

Two regions of the zinc ribbon which are critical for TFIIS function have now been identified, the sequences Asp-Glu and Arg-Trp (Figure 6). Earlier, we proposed that Asp-Glu, together with the RNA polymerase active site, participated in the enhancement of ribonuclease activity. This proposal has gained some support from the recent finding that the RNA polymerase active site includes a nonvariant sequence DFDGD which binds catalytic metal ion (*12*). This scenario is reminiscent of the mechanism proposed for *E. coli* DNA polymerase in which three acidic residues form its active site (*11*). Since pol II ribonuclease activity is extremely slow, it is conceivable that the spatial arrangement of the active site acidic residues (i.e., DFDGD) is not optimal and that the zinc ribbon DE residues of TFIIS create an optimal arrangement by substitution or addition of acidic residues, leading to enhanced ribonuclease activity. The role of the Arg-Trp side chains in ribonuclease activity enhancement may be to localize the TFIIS-modified active center to the disrupted site in the duplex (such as weakly paired and mispaired nucleotides) and modify the site by altering the stereochemistry of the phosphoryl moiety to facilitate rapid hydrolysis. Several lines of investigation are being used to test these proposals.

ACKNOWLEDGMENT

We thank Marsha Rosner, Tao Pan, Terry Martin, Mark Hochstrasser, and Francois Kezdy for critical reading of the manuscript, Somnath Bhattacharyya for preparation of the figures, and Janie Booker for preparation of the manuscript.

REFERENCES

1. Uptain, S., Kane, C., and Chamberlin, M. (1997) *Annu. Rev. Biochem. 66*, 117−172.
2. Reines, D., Conway, J. W., and Conway, R. C. (1996) *Trends Biochem. Sci. 21*, 351−355.
3. Kerppola, T. K., and Kane, C. M. (1988) *Mol. Cell. Biol. 8*, 4389−4394.
4. Agarwal, K., Baek, K., Jeon, C., Miyamoto, K., Ueno, A., and Yoon, H. (1991) *Biochemistry 30*, 7842−7851.
5. Izban, M. G., and Luse, D. S. (1993) *J. Biol. Chem. 268*, 12874−12885.
6. Reines, D. (1992) *J. Biol. Chem. 267*, 3795−3800.
7. Wang, D., and Hawley, D. K. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 843−847.
8. Jeon, C., Yoon, H., and Agarwal, K. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9106−9110.

9. Dombrowski, A. J., Walter, W. A., Record, M. T., Jr., Siegele, D. A., and Gross, C. A. (1992) *Cell 70*, 501−512.
10. Rudd, M. D., Izban, M. G., and Luse, D. S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 8057−8061.
11. Steitz, T. A. (1993) *Curr. Opin. Struct. Biol. 3*, 31−38.
12. Zaychikov, E., Martin, E., Denissova, L., Kozlov, M., Markovtsov, V., Kashlev, M., Heumann, H., Nikiforov, V., Goldfarb, A., and Mustaev, A. (1996) *Science 273*, 107−109.
13. Jeon, C. J. (1993) Ph.D. Thesis, pp 85−92, University of Chicago, Chicago.
14. Yoon, H. S. (1993) Ph.D. Thesis, pp 103−126, University of Chicago, Chicago.
15. Qian, X., Gozani, S. N., Yoon, H. S., Jeon, C. J., Agarwal, K., and Weiss, M. A. (1993) *Biochemistry 32*, 9944−9959.
16. Brun, F., Toulme, J. J., and Héléne, C. (1975) *Biochemistry 14*, 558−563.
17. Qian, X., Jeon, C. J., Yoon, H. S., Agarwal, K., and Weiss, M. A. (1995) *Nature 376*, 279.
18. Gabbay, E. J., Sanford, K., Baxter, C. S., and Kapicak, L. (1973) *Biochemistry 12*, 4021−4029.
19. Toulme, J. J., and Héléne, C. (1976) *J. Biol. Chem. 252*, 244−249.
20. Lakowicz, J. R. (1986) in *Principles of Fluorescence Spectroscopy*, p 347, Plenum Press, New York.
21. Ishida, T., Katsuta, M., Inoue, M., Yamagata, Y., and Tomita, K. (1983) *Biochem. Biophys. Res. Commun. 115*, 849−854.
22. Sha, M., Ferré-D'Amaré, A. R., Burley, S. K., and Goss, D. J. (1995) *J. Biol. Chem. 270*, 19325−19328.
23. Pörschke, D., and Rosenberg, J. (1981) *Biophys. Chem. 13*, 283−290.
24. Werner, M. H., Gronenborn, A. M., and Clore, G. M. (1996) *Science 271*, 778−784.
25. Khamis, M. I., Casas-Finet, J. R., Maki, A. H., Ruvols, P. P., and Chase, J. W. (1987) *Biochemistry 26*, 3347−3354.
26. Kowalczykowski, S. C., Lonberg, N., Newport, J. W., and von Hippel, P. H. (1981) *J. Mol. Biol. 145*, 75−104.
27. Mascotti, D. P., and Lohman, T. M. (1993) *Biochemistry 32*, 10568−10579.
28. Martin, F. H., and Tinoco, I., Jr. (1980) *Nucleic Acids Res. 8*, 2295−2299.
29. Hall, K. B. (1993) *Curr. Opin. Struct. Biol. 3*, 336−339.
30. Ratmeyer, L., Vinayak, R., Zhong, Y. Y., Zon, G., and Wilson, W. D. (1994) *Biochemistry 33*, 5298−5308.
31. Jeon, C. J., and Agarwal, K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13677−13682.
32. Behmoaras, T., Toulmé, J. J., and Héléne, C. (1981) *Nature 292*, 858−859.
33. Valerie, K. (1995) *Nucleic Acids Res. 23*, 3764−3770.
34. Mazumder, A., Gerlt, J. A., Absalon, M. J., Stubbe, J., Cunningham, R. P., Withka, J., and Bolton, P. H. (1991) *Biochemistry 30*, 1119−1126.
35. Blake, R. D., Massoulié, J., and Fresco, J. R. (1967) *J. Mol. Biol. 30*, 291−308.
36. Maher, J. L., Dervan, P. B., and Wold, B. (1992) *Biochemistry 31*, 70−81.

BI980924N