Pergamon

Bioorganic & Medicinal Chemistry 7 (1999) 1075–1079

BIOORGANIC & MEDICINAL CHEMISTRY

Molecular Basis of HIV-1 TAR RNA Specific Recognition by an Acridine *tat*-Antagonist

Nathalie Gelus,[a] François Hamy[b] and Christian Bailly[a,*]

[a]*INSERM Unité 524, IRCL, Place de Verdun, 59045 Lille, France*
[b]*Pharma Research, Novartis Pharma Inc., 4002 Basle, Switzerland*

Received 10 August 1998; accepted 26 October 1998

Abstract—We investigated the interaction of a highly potent acridine-based *tat*-antagonist with the TAR RNA of HIV-1. The wild type TAR RNA and three mutants with U→C23, G·C→C·G26-39 or G·C→A·U26-39 substitutions were used as substrates to study the molecular basis of drug-TAR RNA complex formation. Melting temperature and RNase protection experiments reveal that the G·C26-39 pair is a critical element for specific major groove recognition of TAR at the pyrimidine bulge. The results provide a rational basis for future design of optimized *tat*/TAR inhibitors.

Introduction

The discovery of drugs active against acquired immune deficiency syndrome (AIDS) is one of the most critical challenges of the fall of this century. Significant but limited clinical successes for the treatment of AIDS have been achieved with inhibitors of reverse transcriptase and protease.[1,2] However, the quest for improved anti-HIV-1 drugs remains more pressing than ever because of the growing prevalence of the virus throughout the world and the increasing resistance of the HIV to existing drugs.

The RRE and TAR RNA which bind to the *rev* and *tat* proteins, respectively, are both essential for the regulation of HIV-1 gene expression.[3,4] Therefore, they represent an attractive therapeutic target for the design of new classes of antiviral drugs.[5–7] The past two years have seen remarkable progress in the study of the structure of these two bulged RNA and their function and interaction with small molecules.[8–11] Aminoglycoside antibiotics were found to competitively block the binding of the *rev* protein to its viral recognition site RRE, thereby inhibiting virus proliferation.[12–17] A few aromatic heterocyclic molecules (e.g. diphenylfurans, cyclic polycations) can also interact with the RRE so as to inhibit *rev* binding.[18–20] Concerning TAR, only a small number of *tat*-antagonists have been developed, including a variety of *tat*-derived peptides and peptoids and a few heterocyclic molecules.[21–25]

Very recently, a new class of polyamine-acridine-based compounds was shown to antagonize *tat* binding.[26] The most potent compound in the series, CGP40336A (Figs. 1 and 2), inhibits *tat*-TAR interaction with half-maximal effect at a concentration (C_{50}) of 22 nM whereas concentrations of about 0.5–3 μM are usually required to obtain similar extent of protein binding inhibition with aminoglycoside or other heterocyclic inhibitors.[25] This acridine derivative is one of the most potent *tat*/TAR low molecular weight inhibitors known to date and therefore represents a significant lead for the development of novel anti-HIV drugs.

A hypothetical model, depicted in Figure 1, has been proposed for the binding of CGP40336A into the major groove at the pyrimidine bulge of TAR.[26] On the basis of NMR and footprinting data, it was proposed that the acridine ring stacks over bases U23 and A22. Two potential hydrogen bonds between (i) the methoxy acridine substituent and the exocyclic amino group of C39 and (ii) the amino acridine group and the N7 of G26, would permit specific recognition of the G26·C39 base pair (Fig. 2). In addition, the positively charged polyamine tail attached to the planar chromophore is believed to increase the stability of the drug–TAR complex via salt bridges with the phosphate groups of the RNA.[26]

The rational development of optimized analogues of CGP40336A requires a better understanding of the factors governing the specificity of interaction between the bulge region of TAR and the acridine. In particular, it is essential to determine the influence of the G26·C39 pair (which is essential for *tat* binding)[4] for specific recognition. In this context, we specifically addressed the

Key words: TAR RNA; antiviral drugs; acridine; HIV-1.
*Corresponding author. Tel.: (+33)-320-169218; fax: (+33)-320-169229; e-mail: bailly@lille.inserm.fr

0968-0896/99/$ - see front matter
PII: S0968-0896(99)00030-9

validity of the CGP40336A-TAR model in Figure 1 by comparing the binding of the drug to the wild type TAR RNA and three analogues (Fig. 3). We constructed two TAR mutants for which the target G·C pair is either reversed to a C·G pair or replaced with an A·U pair. A third mutant having the U23 residue replaced with a C residue was also prepared so as to determine the role of this highly conserved uridine residue which is known to be critical for binding of the *tat* protein.[3,4] The interaction of the acridine CPG40336A with these 60 nucleotides long RNA was studied by RNAase A footprinting and melting temperature experiments.

Results and Discussion

Neither the modifications of the G·C pair nor the U→C substitution affect the overall structure of the RNA.

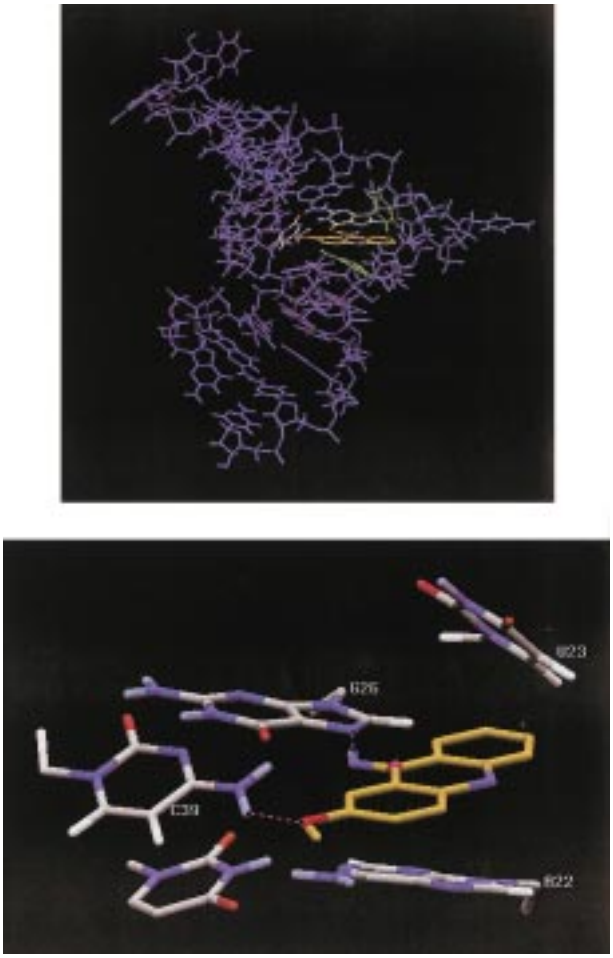

Figure 1. Representation of CGP40336A binding site within the major groove cavity of the TAR RNA. Modeling was performed on an Octane workstation (Silicon Graphics) using the MacroModel software (Dept. of Chemistry, Columbia University). CGP40336A was manually docked into the argininamide-bound TAR RNA for which atomic coordinates were obtained from the SwissProt databank, as described in ref 9. The top panel shows the stacking of the drug over A22 and U23 residues (in green) and the interaction with the adjacent G26·C39 pair. The closer view of the lower panel shows the potential H-bonding interaction between the acridine chromophore of the drug and the G26·C39 pair. For clarity, the polyamine tail attached to the acridine is not shown. This positively charged substituent is believed to interact with the RNA phosphate groups but its exact positioning cannot be determined.

Circular dichroism spectra of wild type and modified TAR superimpose and are characteristic of an A-form conformation. All four CD spectra (not shown) display a weak negative band at 235 nm adjacent to a large positive band centered at 265 nm characteristics of a A-form helix. Moreover, the melting temperatures (T_m) of the different RNA show relatively little variations. T_m of 54.1, 54.3, 51.3 and 53.3 °C were determined for WT, C·G26-39, A·U26-39 and C23 TAR, respectively (average of three independent measurements). In the presence of CGP40336A, the T_m values increase considerably and the ΔT_m values reach 24–25 °C for WT TAR and the mutants C·G26-39 and U23 (Fig. 4). Interestingly, the shift of the T_m is much weaker with A·U26-39 indicating that the replacement of the G·C pair with an A·U pair reduces the capacity of the drug to stabilize the double stranded RNA structure. There is no doubt that the G·C pair is important for the recognition of TAR by the acridine compound. RNase protection experiments fully support this view.

The cleavage of the wild type TAR RNA by RNase A is largely affected by CGP40336A. The strong cutting site

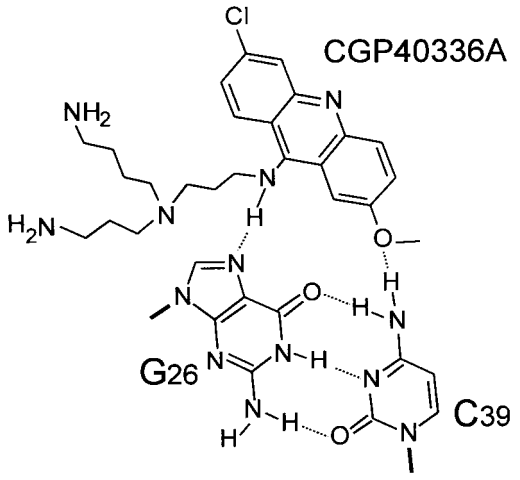

Figure 2. Structure of CGP40336A and its interaction with the G26·C39 base pair. The potential H-bonds between acridine NH and N7 guanine as well as between acridine methoxy group and cytosine amino group are indicated (dotted lines).

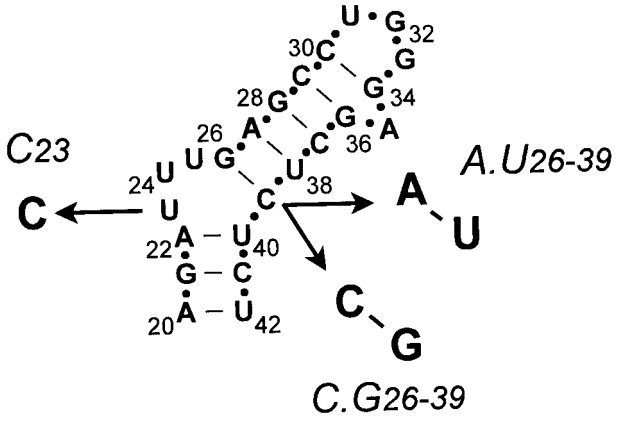

Figure 3. Sequence and secondary structure of the bulge and loop regions of TAR. The positions of the base modifications corresponding to the three mutants are indicated.

at the UCUC sequence facing the pyrimidine bulge is completely abolished in the presence of the acridine, even at a concentration as low as 0.5 μM. Titrations indicated that a concentration of 0.15 μM was sufficient to reduce the cleavage of WT TAR by 50%. With the mutants, cleavage inhibition at the UCUC site can also be detected but the extent of protection is reduced. From the typical footprinting gels shown in Figure 5 and several others, we determined the minimum drug concentration required for complete inhibition of RNase A cleavage (Fig. 6). The substitution of an A·U pair for the G·C pair profoundly reduces the capacity of the drug to bind to the TAR RNA. A concentration of ~7 μM is needed to reduce RNase A cutting of A·U26-39 TAR by 50%. This concentration is 46-times higher than that used with the wild type RNA to observe similar extent of nuclease protection. The two other mutations, G·C→C·G reversion and U→C substitution, are better tolerated. The uridine residue may contribute partially to the drug–RNA recognition process. With U23 TAR, the binding is only reduced by a factor of about 11 compared to the wild type sequence and with C·G26-39 TAR the RNase protection effect decreased by about 14-fold, which is significant but much less than with A·U26-39 TAR.

The capacity of the highly potent *tat*-antagonist CGP40336A to interact with the TAR RNA largely depends on the formation of specific contacts with the G26·C39 pair which delimits the *tat* binding major groove cavity. The T_m and RNase protection data reported here support the model presented in Figure 1. The drug may engage H-bonds with C·G26-39 TAR as with WT TAR providing that the acridine chromophore is rotated by 180°. However, in this case the bulky chloro group (now stacking over A22 residue) must hinder the correct anchorage of the drug into the major groove pocket. As a result of the steric hindrance, the extent of RNase protection is markedly decreased. As regards the reduced binding of the drug to C23 TAR, the effect is most likely attributable to the newly introduced bulky amino group on the cytosine residue which

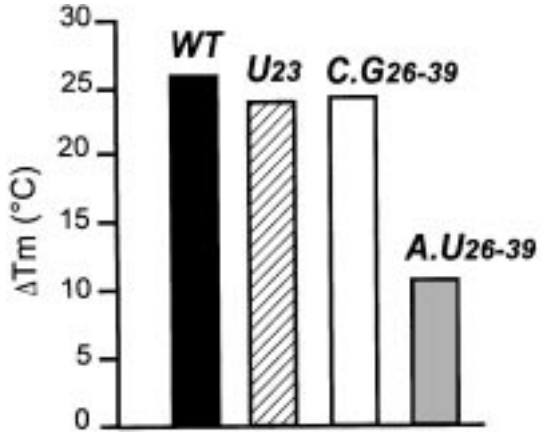

Figure 4. Variation of the melting temperatures (ΔT_m) for the four RNA (10 μM each) in the presence of 0.2 μM CGP40336A. T_m measurements were performed in BPE buffer pH 7.1 (6 mM Na_2HPO_4, 2 mM $Na_2H_2PO_4$, 1 mM EDTA) using 10 μM RNA in 1 mL quartz cuvettes at 260 nM with a heating rate of 1 °C/min.

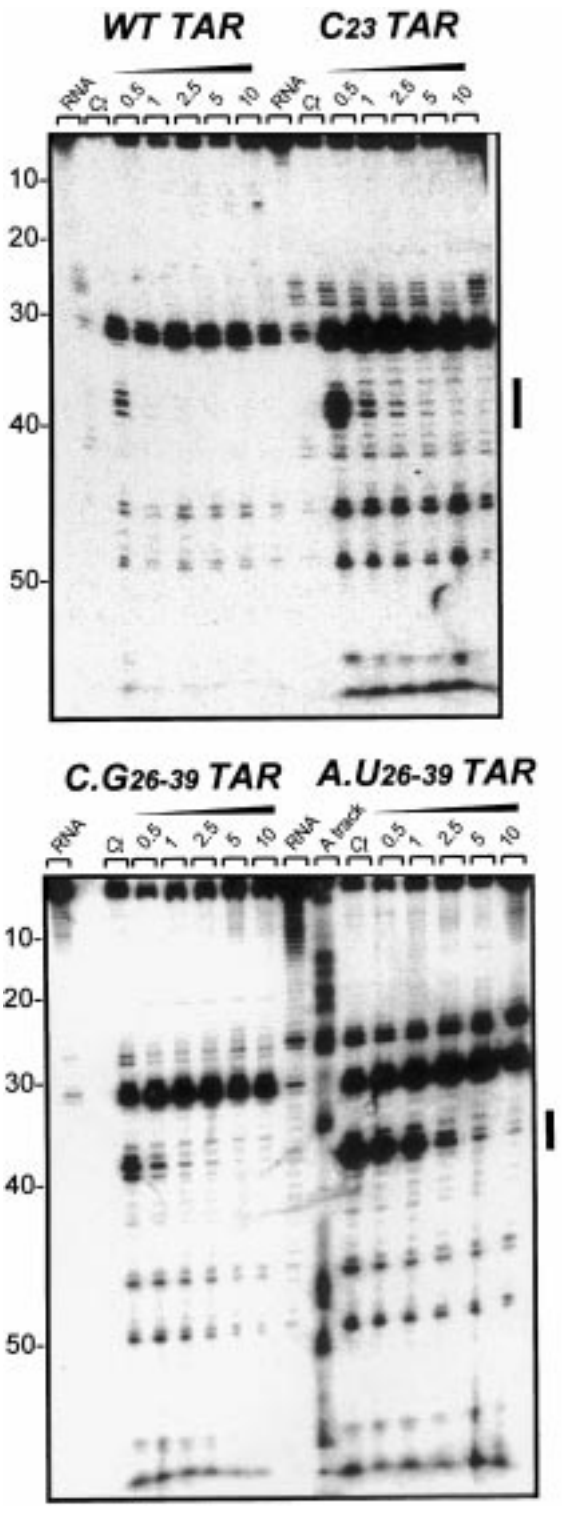

Figure 5. Phosphorimages showing RNAase A footprinting of CGP40336A on the wild type TAR RNA and the mutants C·G26-39, A·U26-39 and C23 TAR. In each case, the RNA was 3′-end labelled with [^{32}P]cytidine biphosphate and T4 RNA ligase. The cleavage products of RNAase digestion were resolved on a 10% polyacrylamide gel containing 8 M urea. The concentration (μM) of the drug is shown at the top of the appropriate gel lanes. Control tracks labelled 'Ct' contained no drug. Track labelled A represents diethylpyrocarbonate-aniline markers specific for adenines. Lane marked 'RNA' contain the [^{32}P]labelled RNA alone, incubated without drug or enzyme; this sample serves as a control to assess background nicking of the RNA preparation. Numbers at the left side of the gel refer to the numbering scheme used in Figure 3.

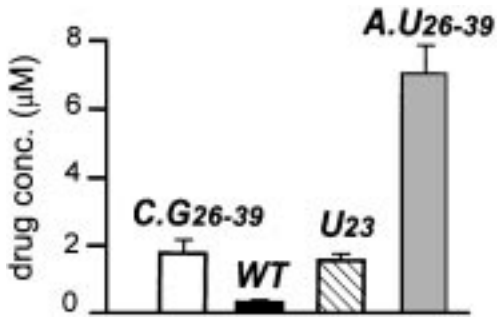

Figure 6. Minimum drug concentration (μM) required to completely inhibit RNAase cleavage at the UCUC binding site in the different RNA. Data are compiled from quantitive analysis of four series of phosphor-images such as those shown in Figure 5 and must be considered as a set of averaged values. The specific concentrations were obtained from plots of protection versus drug concentration. The differences in sample loading between the different gels were taken into account.

projects in the major groove so as to interfere with the access of the drug. The U→C substitution not only replaces an acceptor C=O with a donor group C-NH_2 at the 4 position but also removes a hydrogen-bond donor at N^3-H. It is therefore logical that stacking interactions are affected. With A·U26-39 TAR, the H-bond between NH acridine and N7 purine (A instead of G) can be maintained whereas the H-bond involving the OCH_3 acridine cannot form because of the loss of the exocyclic NH_2 pyrimidine. One would anticipate that the loss of one of the two drug–RNA H-bonds should correspond to a drastic reduction of the drug–RNA interaction and this is exactly what we observed.

Conclusion

The results show that CGP40336A can discriminate between C·G or A·U pairs at position 26-39 as opposed to the natural G·C pair. The demonstration that this G·C pair is a major recognition element for CGP40336A–TAR complex formation opens exciting perspectives for the development of specific analogues. Thus the results provide a unique molecular basis for the design of TAR-targeted anti-HIV drugs.

Experimental

Chemicals

A pure sample of CGP40336A was provided by Novartis. The synthesis of the drug has been reported.[26] All other chemicals were analytical grade reagents, and solutions were prepared with doubly distilled sterile water to prevent from nuclease contamination. Tubes and tips were treated with 1% diethylpyrocarbonate (DEPC from Sigma).

In vitro transcription

Synthetic oligonucleotides corresponding to the wild type or mutated TAR sequences were cloned between *Hind*III and *Eco*RI sites of the pUC19 plasmid. After digestion with *Eco*RI, the RNA was transcribed as a run-off product of 60 nucleotides from the T3 RNA polymerase promoter. In each case the transcript includes an additional G residue on the 3′-end derived from the *Eco*RI cleavage site. Transcription reaction was performed in buffer containing 40 mM Tris–HCl, pH 7.4, 25 mM NaCl, 16 mM $MgCl_2$, 10 mM DTT and 1 mM NTPs. The reaction was initiated by addition of 10 μg linearized plasmid DNA template and 40 μg T3 RNA polymerase and incubated for 2 h at 37 °C. Nucleic acids were then fractionated on a 10% (w/v) polyacrylamide gel containing 8 M urea in TBE buffer (89 mM tris-borate pH 8.3, 10 mM EDTA). After electrophoresis, the RNA was eluted in water for 18 h at 4 °C and precipitated with ethanol. The RNA was resuspended in water to give a 500 μM stock solution (ε^{260}/phosphate = 10688 $M^{-1}\times cm^{-1}$). For the footprinting experiments, the RNA was 3′-end labelled with [^{32}P]cytidine biphosphate and T4 RNA ligase and then repurified from a 10% denaturing acrylamide gel.

Absorption spectroscopy and melting temperature studies

Absorption spectra were recorded on a Uvikon 943 spectrophotometer. The 12-cells holder was thermostated with a Neslab RTE 111 cryostat. Drug–RNA complexes were prepared by adding aliquots of a concentrated CGP40336A solution to a RNA solution at constant concentration (usually 20 μM) in BPE buffer pH 7.1 (6 mM Na_2HPO_4, 2 mM $Na_2H_2PO_4$, 1 mM EDTA). A heating rate of 1 °C/min was used and data points were collected every 30 s. The temperature inside the cuvette was monitored by using a thermocouple. The absorbance at 260 nM was measured over the range 25–90 °C in 1 cm path length reduced volume quartz cells. The 'melting' temperature T_m was taken as the mid-point of the hyperchromic transition determined from first derivatives plots. The reproducibility of the T_m measurements is ±1 °C.

RNAase A footprinting, gel electrophoresis and data processing

The procedure for the footprinting experiments has been described previously.[27] Briefly, samples of the labelled RNA fragment were incubated with a buffered solution containing the desired drug concentration. After 20 min incubation at 4 °C to ensure equilibration, the digestion was initiated by addition of the RNAase A solution. After 1 min incubation at room temperature, the reaction was stopped by freeze drying and samples were lyophilized. The RNA in each tube was resuspended in 5 μL of formamide-TBE loading buffer, denatured at 90 °C for 4 min then chilled in ice for 4 min prior to loading on to a 0.3 mM thick, 10% polyacrylamide gel containing 8 M urea and TBE buffer (89 mM Tris base, 89 mM boric acid, 2.5 mM Na_2 EDTA, pH 8.3). After 2 h electrophoresis at 1500 V, the gel was soaked in 10% acetic acid for 10 min, transferred to Whatman 3MM paper, dried under vacuum at 80 °C and then analyzed on the phosphorimager (Molecular Dynamics). Each resolved band on the autoradiograph was assigned to a particular bond within the RNA fragment by comparison of its position

relative to sequencing standards generated by treatment of the RNA with diethylpyrocarbonate followed by aniline-induced cleavage at the modified bases (A track).

Acknowledgements

This work was done under the support of a research grant (to C.B.) from the Agence Nationale de la Recherche sur le SIDA (ANRS). N.G. is currently supported by a fellowship from the Institut de Recherches sur le Cancer de Lille.

References

1. Vasudevachari, M. B.; Battista, C.; Lane, H. C.; Psallidopoulos, M. C.; Zhao, B.; Cook, J.; Palmer, J. R.; Romero, D. L.; Tarpley, W. G.; Salzman, N. P. *Virology* **1992**, *190*, 269.
2. Hammer, S. M. *AIDS* **1996**, *10*, S1.
3. Gait, M. J.; Karn, J. *Trends Biochem. Sci.* **1993**, *18*, 255.
4. Churcher, M. J.; Lamont, C.; Hamy, F.; Dingwall, C.; Green, S. M.; Lowe, A. D.; Butler, P. J. G.; Gait, M. J.; Karn, J. *J. Mol. Biol.* **1993**, *230*, 90.
5. Fernandez-Saiz, M.; Schneider, H.-J.; Sartorius, J.; Wilson, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 4739.
6. Chen, Q.; Shafer, R. H.; Kuntz, I. D. *Biochemistry* **1997**, *36*, 11402.
7. Pearson, N. D.; Prescott, C. D. *Chem. Biol.* **1997**, *4*, 409.
8. Berkhout, B. *Prog. Nucleic Acid Res. Mol. Biol.* **1996**, *54*, 1.
9. Aboul-ela, F.; Karn, J.; Varani, G. *J. Mol. Biol.* **1995**, *253*, 313.
10. Aboul-ela, F.; Karn, J.; Varani, G. *Nucleic Acids Res.* **1996**, *24*, 3974.
11. Chow, C. S.; Bogdan, F. M. *Chem. Rev.* **1997**, *97*, 1489.
12. Zapp, M. L.; Stern, S.; Green, M. R. *Cell* **1993**, *74*, 969.
13. Werstuck, G.; Zapp, M. L.; Green, M. R. *Chem. Biol.* **1996**, *4*, 129.
14. Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 3641.
15. Wallis, M. G.; Schroeder, R. *Prog. Biophys. Molec. Biol.* **1997**, *67*, 141.
16. Wang, H.; Tor, Y. *J. Am. Chem. Soc.* **1997**, *119*, 8734.
17. Wang, H.; Tor, Y. *Angew. Chem., Int. Ed.* **1998**, *37*, 109.
18. Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R. *Bioorg. Med. Chem.* **1997**, *5*, 1149.
19. Ratmeyer, L.; Zapp, M. L.; Green, M. R.; Vinayak, R.; Kumar, A.; Boykin, D. W.; Wilson, W. D. *Biochemistry* **1996**, *35*, 13689.
20. Li, K.; Fernandez-Saiz, M.; Rigl, C. T.; Kumar, A.; Ragunathan, K. G.; McConnaughie, A. W.; Boykin, D. W.; Schneider, H.-J.; Wilson, W. D. *Bioorg. Med. Chem.* **1997**, *5*, 1157.
21. Weeks, K. M.; Crothers, D. M. *Cell* **1991**, *66*, 577.
22. Ratmeyer, L.; Vinayak, R.; Zon, G.; Wilson, W. D. *J. Med. Chem.* **1992**, *35*, 966.
23. Hamy, F.; Felder, E. R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3548.
24. Huq, I.; Wang, X.; Rana, T. M. *Nature Struct. Biol.* **1997**, *4*, 881.
25. Mei, H.-Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. *Bioorg. Med. Chem.* **1997**, *5*, 1173.
26. Hamy, F.; Brondani, V.; Flörsheimer, A.; Stark, W.; Blommers, M. J. J.; Klimkait, T. *Biochemistry* **1998**, *37*, 5086.
27. Bailly, C.; Colson, P.; Houssier, C.; Hamy, F. *Nucleic Acids Res.* **1996**, *24*, 1460.