Bull. Chem. Soc. Jpn., 74, 1139—1143 (2001)

Conformational Study of the Tandem Repeat Sequence in RNA Polymerase II by Circular Dichroism Spectroscopy

Isao Ohiso, Masahiko Tsunemi,† Budiao Zhao, and Norio Nishi*

Graduate School of Environmental Earth Science, Hokkaido University, Kita-ku, Sapporo 060-0810

†Peptide Institute, Inc., Ina, Minoh-shi, Osaka 562-8686

(Received December 4, 2000)

The largest subunit of RNA polymerase II has an interesting C-terminal domain, which consists of a multiple tandem repeat of a heptapeptide with the consensus sequence Ser–Pro–Thr–Ser–Pro–Ser–Tyr. However, the functional role of this sequence is unclear. One might assume that the conformation of the tandem repeat is very important for its function. The conformation of poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) and H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH, corresponding to the repeat and the repeating unit of the C-terminal domain, respectively, were investigated exhaustively by circular dichroism (CD). The CD spectrum of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH indicated the presence of a turn structure in water, which was further stabilized in 2,2,2-trifluoroethanol (TFE) and in acetonitrile. The CD spectra of the polyheptapeptide poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) suggested that the disordered conformation was predominant in water and that the turn structure was stabilized by increasing the content of TFE or acetonitrile in the solvent. We suggest that the polyheptapeptide forms not only a turn structure at the heptapeptide unit Ser–Pro–Thr–Ser–Pro–Ser–Tyr, but also a kind of super conformation induced by the periodicity of the sequence in TFE or acetonitrile.

RNA polymerases are enzymes known to catalyze RNA synthesis from a DNA template strand. Procaryotes have a single RNA polymerase to synthesize all RNAs, while eucaryotes have three different RNA polymerases: RNA polymerase I, II, and III. Among them, RNA polymerase II transcribes the genes that encode for proteins. On the other hand, RNA polymerase I and III synthesize the large ribosomal RNAs and various small RNAs, respectively. It is known that the RNA polymerase II complex consists of two large subunits (molecular weight > 100,000) and 7–12 smaller subunits, and that the combined molecular weight of the complex exceeds 500,000.[1] The largest subunit of RNA polymerase II has a very interesting C-terminus consisting of multiple tandem repeat sequences of seven amino acid residues having the consensus Ser–Pro–Thr–Ser–Pro–Ser–Tyr. The number of tandem repeats can vary; for example, there are 26 units in yeast (*S. cerevisiae*),[2] 44 units in *D. melanogaster*,[3] and 52 units in hamster.[4] The largest subunit of RNA polymerase II is homologous with the largest subunits of RNA polymerase I and III.[2] However, the tandem repeat sequence is not found in the largest subunits of eucaryotic RNA polymerase I[5] and III,[2] or in the procaryotic RNA polymerase.[6] It was shown by deletion analysis that approximately half of the repeating units of this tandem repeat are necessary for the function of this enzyme.[3,4,7,8] Though the tandem repeat plays an essential role in the initiation step of transcription,[2–4,7–12] its functional role is unclear. It has been suggested that the tandem repeat binds to transcription factors to form the transcription pre-initiation complex.[4,8,12–15] On the other hand, it is interesting that the tandem repeat contains two copies of the SPXX motif. The SPXX motif is found in various DNA-binding proteins, such as histones and gene regulatory proteins, and has been suggested to be a DNA-binding motif.[16] By examining the effect of synthetic peptides containing one and two units of the consensus sequence Tyr–Ser–Pro–Thr–Ser–Pro–Ser–Tyr and Tyr–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–Ser–Pro–Thr–Ser–Pro–Ser–Tyr, respectively,[17] Suzuki suggested that the heptapeptide repeat might bind to DNA through the intercalation of its tyrosyl side-chain.

The heptapeptide sequence consists of proline residues and three kinds of residues having side-chain hydroxy groups. Proline is a peculiar amino acid in that it is conformationally restricted due to the limited dihedral angle ϕ of its backbone and steric interactions between its δ carbon and the preceding residue because of the pyrrolidine ring formed with the side-chain. Moreover, the proline residue cannot be a hydrogen bond donor because the amide proton is replaced by a CH_2 group. The tandem repeat in RNA polymerase II is estimated to have a characteristic conformation different from ordinary conformations, such as α-helix and β-sheet, because it is not only proline-rich, but also a periodic sequence. Its characteristic conformation may play an important role in the function of the C-terminal domain. A detailed investigation using a polypeptide corresponding to the tandem repeat is important to clarify the functional role of this repeat. In our previous study,[18] the sequential polyheptapeptide corresponding to this tandem repeat, poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr), and the heptapeptide corresponding to the repeating unit, H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH, were synthesized.

In the present study, the conformation of the polyheptapeptide poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) was investigated exhaustively by circular dichroism and compared with that of the monomeric heptapeptide, H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–

OH.

Experimental

The syntheses of the heptapeptide corresponding to the unit of the tandem repeat H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH and the polyheptapeptide corresponding to the tandem repeat poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) were previously reported.[18] The molecular weight of the polyheptapeptide was estimated by SDS-PAGE to be 60,000–70,000 (83–97 heptapeptide units). All circular dichroism (CD) spectroscopy was performed on a Jasco 720A spectropolarimeter. The CD spectra were collected using a quartz cell of 1.0 mm path length at 20 °C. The concentration of the heptapeptide and polyheptapeptide was approximately 0.1 mg mL^{-1}. All spectra were expressed as mean residue ellipticity, [θ] (deg·cm^2/dmol).

Results

CD Spectra in Water. The CD spectra of the monomeric heptapeptide H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in water and in NaCl solution are shown in Fig. 1a. The CD spectrum in water showed a negative maximum at 208 nm and a negative minimum at 199 nm. The CD spectra of the polymeric heptapeptide poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in water and in NaCl solution are shown in Fig. 1b. The CD spectrum in water showed only a negative maximum at 202 nm. The spectra of both the heptapeptide and the polyheptapeptide showed an increase in the strength of ellipticity in a NaCl solution. The CD spectra of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH and poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in a 100 mM NaCl solution at various pH levels are shown in Fig. 2. Changes in the pH induced only small changes in the spectra. A positive maximum near 245 nm appeared on the CD spectra of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in the high pH solution, but it was not found in the spectra of poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr). This CD band is attributable to the contribution of phenolic anions on the dissociated tyrosyl side-chain, but its conformational interpretation remains unclear.

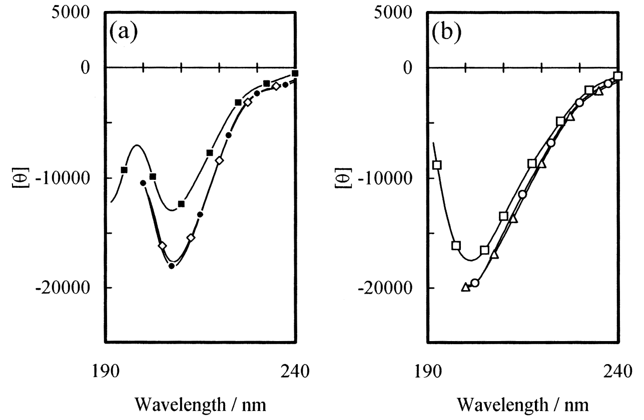

Fig. 1. CD spectra of the model peptides at various NaCl concentrations. (a) H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in water (—■—), in 50 mM NaCl (—●—), and in 150 mM NaCl (—◇—); (b) poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in water (—□—), in 50 mM NaCl (—○—), and in 150 mM NaCl (—△—).

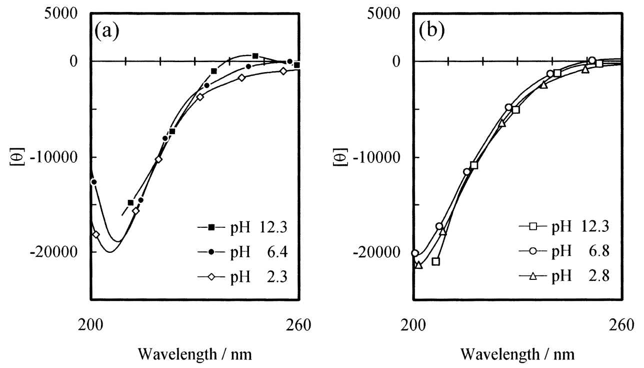

Fig. 2. CD spectra of the model peptides at various pH values in 100 mM NaCl; (a) H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH, (b) poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr).

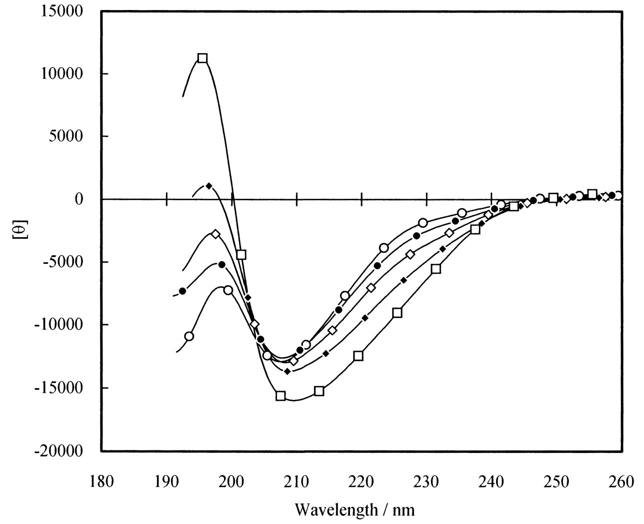

Fig. 3. CD spectra of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in TFE/water. (—○—, TFE 0%; —●—, TFE 25%; —◇—, TFE 50%; —◆—, TFE 75%; —□—, TFE 100%)

The CD Spectra in Trifluoroethanol/Water. The CD spectra of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in water containing varying amounts of 2,2,2-trifluoroethanol (TFE) are shown in Fig. 3. In water, the spectrum had a negative maximum at 208 nm and a negative minimum at 199 nm. By increasing the content of TFE in the TFE/water mixture, these CD bands increased in intensity, exhibiting an isodichroic point at 204 nm. The CD spectrum in TFE showed a negative maximum at 210 nm and a positive maximum at 196 nm. The isodichroic point is interpretable as an observation of the conformational change between two components. The CD spectra of poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in TFE/water are shown in Fig. 4. The negative maximum at 202 nm of the spectrum in water vanished upon increasing the content of TFE in the TFE/water mixture without showing an isodichroic point. In TFE, a negative maximum at 207 nm and a negative minimum at 200 nm were observed.

The CD Spectra in Acetonitrile/Water. Both H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH and poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) are insoluble in acetonitrile alone. The CD spec-

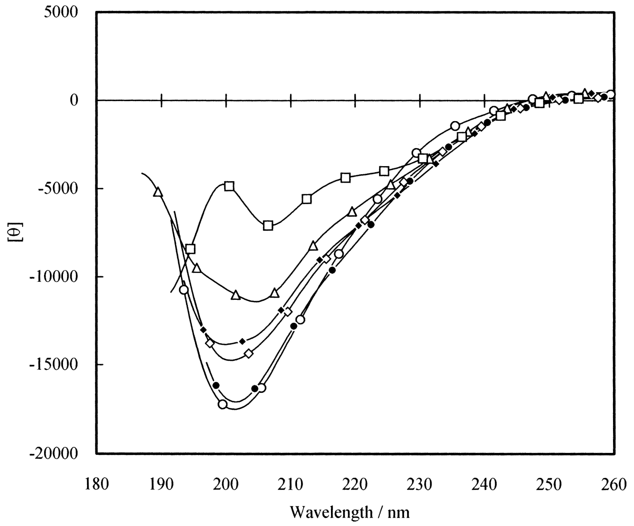

Fig. 4. CD spectra of poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in TFE/water. (—○—, TFE 0%; —●—, TFE 25%; —◇—, TFE 50%; —◆—, TFE 75%; —△—, TFE 87.5%; —□—, TFE 100%)

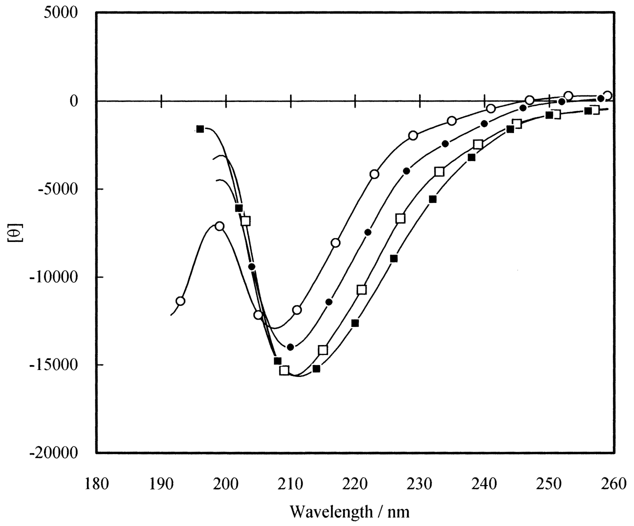

Fig. 5. CD spectra of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in acetonitorile/water. (—○—, acetonitrile 0%; —●—, acetonitrile 50%; —□—, acetonitrile 70%; —■—, acetonitrile 90%)

tra of H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in water containing various amounts of acetonitrile are shown in Fig. 5. As the content of acetonitrile increased, the CD spectrum showed a similar change, as in the case of the TFE/water mixture with an isodichroic point at 205 nm. A negative maximum at 211 nm and a negative minimum at 197 nm were detected in 90% acetonitrile. The CD spectra of poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in acetonitrile/water are shown in Fig. 6. By increasing the content of acetonitrile in the acetonitrile/water mixture, the negative maximum of the spectra in water vanished and a spectral change similar to that found in the TFE/water system was observed with an isodichroic point at 222 nm. A negative maximum at 208 nm appeared in 90% acetonitrile.

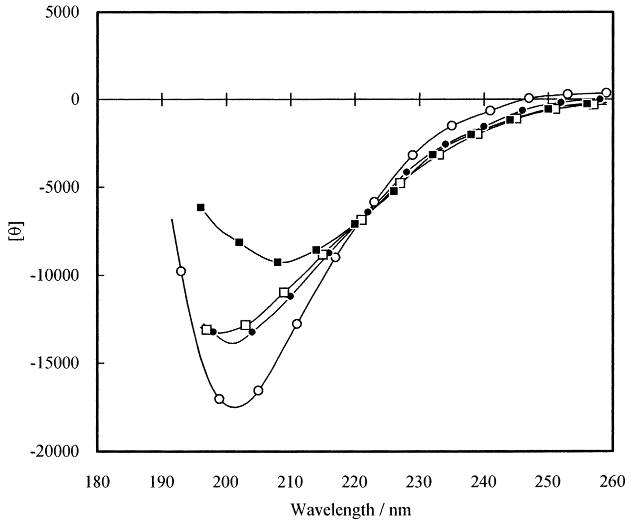

Fig. 6. CD spectra of poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in acetonitrile/water. (—○—, acetonitrile 0%; —●—, acetonitrile 50%; —□—, acetonitrile 70%; —■—, acetonitrile 90%)

Discussion

By changing the NaCl concentration in water for both the monomeric heptapeptide H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH and the polyheptapeptide poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr), the strength of ellipticity of the CD spectra was increased, but the wavelength of the CD bands did not change. This suggests that the NaCl concentration did not influence the conformation of the heptapeptide and the polyheptapeptide. In high-pH solutions, the CD band of the dissociated tyrosyl side-chain appeared for the heptapeptide, but not for the polyheptapeptide. The relationship between the CD band and the conformational property of the tyrosyl side-chain is unknown. In a wavelength region providing information about the backbone conformation, changes in pH did not result in significant spectral changes for the heptapeptide and the polyheptapeptide.

For both the monomeric heptapeptide, H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH, and the polyheptapeptide, poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr), similar changes in the CD spectra were induced by TFE and acetonitrile. In hydrogen-bond formation, water has both proton donor and acceptor capacities, whereas TFE has a strong proton donor capacity and the acetonitrile solvent is a weak hydrogen bond acceptor. Since TFE and acetonitrile are less effective than water in destroying the ordered conformation of proteins or polypeptides, the conformation observed in these solvents may reflect their native conformation. The spectral changes shown in Figs. 3–6 may reflect conformational changes induced by the formation of intramolecular hydrogen bonds favored by these solvents.

To estimate the contents of secondary structures from the measured CD spectra, a curve-fitting analysis was performed. Assuming that a measured CD spectrum, f_m, is a weighted sum of the CD spectra of α-helix, β-sheet, β-turns, and disordered conformation, the spectrum can be expressed as

$$f_m = C_\alpha g_\alpha + C_\beta g_\beta + C_t g_t + C_d g_d, \qquad (1)$$

where g_α, g_β, g_t, and g_d are reference CD spectra for α-helix, β-sheet, β-turns, and disordered conformation, respectively, and C_α, C_β, C_t, and C_d are their weighted coefficients. We used the reference CD spectra proposed by Bolotina et al.,[19,20] which were obtained from experimental CD spectra of five standard proteins whose secondary structures were determined by X-ray crystallography. Based on a least-squares fit, we calculated the coefficients for the CD spectra of the heptapeptide and the polyheptapeptide measured in water and TFE. However, the CD spectra synthesized from the calculated coefficients deviated considerably from the measured CD spectra. The root-mean-square errors (deg cm^2 dmol^{-1}) of the synthesized spectra from the measured spectra of the heptapeptide in water and in TFE were 2.36×10^3 and 1.79×10^3, respectively, and those of the polyheptapeptide in water and in TFE were 3.34×10^3 and 8.91×10^2, respectively. This method is deficient in that only one spectrum is provided as the reference CD spectrum for a β-turn, which is interpreted as the sum of type I, II, and III β-turns minus the sum of type I′, II′, and III′ β-turns. The CD spectra of β-turns calculated theoretically by Woody[21] were very different for each class. The deviated results of the synthesized spectra ruled out the possibility that these experimental spectra appeared accidentally by some combination of three components: α-helix, β-sheet, and a disordered conformation lacking β-turns. Therefore, it was suggested that in water and TFE the heptapeptide and the polyheptapeptide formed some conformations distinct from α-helix, β-sheet, or disordered conformations.

The CD spectra of the heptapeptide and the polyheptapeptide in water and TFE are shown in Fig. 7. We suggest that a negative maximum at 210 nm and a positive maximum at 196 nm of the heptapeptide in TFE are contributions of the component spectrum 1 (Fig. 12A of the reference), which was proposed by Perczel et al. as indicating the presence of a turn structure based on quantitative circular dichroism, NOE, and X-ray data.[22] It was suggested that the heptapeptide in water included the contribution of the component spectrum 1 caused by the turn structure, and that as the content of TFE increased its contribution became more predominant, because TFE favored the turn structure. A similar interpretation might explain the spectrum changes of the heptapeptide in the acetonitrile/water system. These results do not contradict the NMR studies by Harding,[23] who proposed the presence of turn structures in equilibrium with unfolded forms of the octapeptide Tyr–Ser–Pro–Thr–Ser–Pro–Ser–Tyr in water.

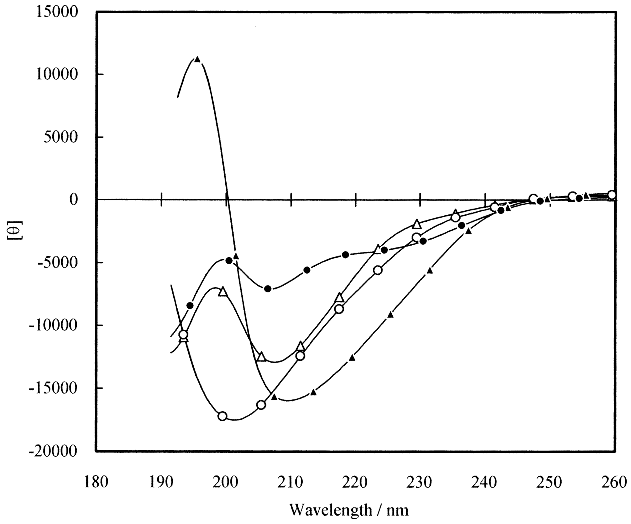

Fig. 7. CD spectra of a model peptide and polypeptide corresponding to the tandem repeat sequence in RNA polymerase II, H–Ser–Pro–Thr–Ser–Pro–Ser–Tyr–OH in water (—△—), and in TFE (—▲—), and poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) in water (—○—) and in TFE (—●—).

On the CD spectra of the polyheptapeptide, the negative maximum at 202 nm in water suggested a predominant disordered conformation in water. The spectral changes resulting from increased concentrations of TFE did not show an isodichroic point, suggesting complicated changes between three or more spectral components. As in the case of the heptapeptide, the contribution of the component spectrum 1 was observed in TFE as a negative maximum at 207 nm and a negative minimum at 200 nm for the polyheptapeptide. Also, in the acetonitrile/water system, it was observed that the contribution of the component spectrum 1 became dominant by increasing the content of acetonitrile, resulting in spectrum changes similar to those observed in the TFE/water system. The only difference was that the isodichroic point was observed in this case. It was suggested that the presence of either TFE or acetonitrile favored the turn structure in the polyheptapeptide in a manner similar to that of the heptapeptide. We propose that the differences between the CD spectra of the heptapeptide and the polyheptapeptide in TFE or 90% acetonitrile arise from the conformational characteristics of the polyheptapeptide caused by multiple tandem repeats of the heptapeptide sequence. It is suggested that the polyheptapeptide includes not only the turn structure caused by the heptapeptide sequence Ser–Pro–Thr–Ser–Pro–Ser–Tyr, but also a kind of super conformation caused by intramolecular hydrogen bond formation within the tandemly repeating sequence in some environments other than water. It is possible that the conformation of the tandem repeat plays a key role in its function. The CD and NMR analysis reported by our group Imai et al.[24] suggested that the polyheptapeptide poly(Ser–Pro–Thr–Ser–Pro–Ser–Tyr) having a super conformation interacts with DNA in the TFE/water system.

These results suggest that the turn structure is a favorable conformation for the consensus heptapeptide sequence Ser–Pro–Thr–Ser–Pro–Ser–Tyr in the tandem repeat of the RNA polymerase II. Moreover, these results suggest that the repeat domain in RNA polymerase II contains a kind of super conformation that includes the turn structure in its native state.

References

1 M. K. Lewis and R. R. Burgess, in "The Enzymes," ed by P. D. Boyer, Academic Press, New York (1982), Vol. 15, pp. 109–153.

2 L. A. Allison, M. Moyle, M. Shales, and C. J. Ingles, *Cell*, **42**, 599 (1985).

3 W. A. Zehring, J. M. Lee, J. R. Weeks, R. S. Jokerst, and A. L. Greenleaf, *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 3698 (1988).

4 L. A. Allison, J. K. C. Wong, V. D. Fitzpatrick, M. Moyle,

and C. J. Ingles, *Mol. Cell. Biol.*, **8**, 321 (1988).

5 J.-M. Buhler, M. Riva, C. Mann, P. Thuriax, S. Memet, J. Y. Micouin, I. Treich, S. Mariotte, and A. Sentenac, in "RNA polymerase and Regulation of Transcription," ed by W. S. Reznikoff, R. R. Burgess, J. E. Dahlberg, C. A. Gross, M. T. Record, Jr., and M. P. Wickens, Elsevier, New York (1987), pp. 25–36.

6 Y. A. Ovchinnikov, G. S. Monastyrskaya, V. V. Gubanov, S. O. Guryev, I. S. Salomatina, T. M. Shuvaeva, V. M. Lipkin, and E. D. Sverdlov, *Nucleic Acids Res.*, **10**, 4035 (1982).

7 M. S. Bartolomei, N. F. Halden, C. R. Cullen, and J. L. Corden, *Mol. Cell. Biol.*, **8**, 330 (1988).

8 M. Nonet, D. Sweetser, and R. A. Young, *Cell*, **50**, 909 (1987).

9 D. Cadena and M. E. Dahmus, *J. Biol. Chem.*, **262**, 12468 (1987).

10 J. L. Christmann and M. E. Dahmus, *J. Biol. Chem.*, **256**, 11798 (1981).

11 N. E. Thompson, T. H. Steinberg, D. B. Aronson, and R. R. Bourgess, *J. Biol. Chem.*, **264**, 11511 (1989).

12 M. Moyle, J. S. Lee, W. F. Anderson, and C. J. Ingles, *Mol. Cell. Biol.*, **9**, 5750 (1989).

13 P. B. Sigler, *Nature*, **333**, 210 (1988).

14 A. Usheva, E. Maldonado, A. Goldring, H. Lu, C. Houbavi, D. Reinberg, and Y. Aloni, *Cell*, **69**, 871 (1992).

15 A. J. Koleske, S. Buratowski, M. Nonet, and R. A. Young, *Cell*, **69**, 883 (1992).

16 M. Suzuki, *J. Mol. Biol.*, **207**, 61 (1989).

17 M. Suzuki, *Nature*, **344**, 562 (1990).

18 N. Nishi, I. Ohiso, N. Sakairi, S. Tokura, M. Tsunemi, and M. Oka, *Biochem. Biophys. Res. Commun.*, **206**, 981 (1995).

19 I. A. Bolotina, V. O. Chekhov, V. Yu. Lugauskas, A. V. Finkel'shtein, and O. B. Ptitsyn, *Mol. Biol. (Eng. Transl.)*, **14**, 701 (1980).

20 I. A. Bolotina, V. O. Chekhov, V. Yu. Lugauskas, and O. B. Ptitsyn, *Mol. Biol. (Eng. Transl.)*, **14**, 709 (1980).

21 R. W. Woody, in "Peptides, Polypeptides and Proteins," ed by E. R. Blout, F. A. Bovey, M. Goodman, and N. Lotan, Wiley, New York (1974), pp. 338–350.

22 A. Perczel, M. Hollósi, B. M. Foxman, and G. D. Fasman, *J. Am. Chem. Soc.*, **113**, 9772 (1991).

23 M. M. Harding, *J. Med. Chem.*, **35**, 4658 (1992).

24 K. Imai, I. Ohiso, T. Hiraoki, M. Tsunemi, M. Oka, S. Tokura, and N. Nishi, *Peptide Chem.*, **1996**, 257 (1997).