Conformation of the Glycopeptide Linkage in Asparagine-Linked Glycoproteins†

C. Allen Bush,* Anthony Duben,‡ and Suresh Ralapati

ABSTRACT: The conformation of the glycopeptide linkage region of the serum-type glycoproteins has been studied by CD and NMR experiments on the linkage compound 2-acetamido-1-*N*-(4-L-aspartyl)-2-deoxy-β-D-glycopyranosylamine (I) and on the model compound 1-*N*-acetyl-2-deoxy-2-acetamido-β-D-glycopyranosylamine (II). The circular dichroism spectra of I and II have been measured in the 170–220-nm region by using a newly constructed vacuum ultraviolet CD instrument. The CD data on these compounds are interpreted in terms of exciton-type coupling between the vicinal acyl amino chromophores. A detailed theory of this coupling is used to relate the measured CD spectra to the dihedral angles τ_1 and τ_2 describing rotation about the exocyclic amide bonds at carbons 1 and 2. Proton NMR experiments have been used to measure the vicinal coupling constants of the amide protons. The amide dihedral angles, determined from these data by means of a Karplus relation commonly used in peptide NMR studies, are consistent with those determined by CD experiments. Vicinal H–H coupling constants of the side-chain protons give information on the asparagine side chain dihedral angle χ^1. On the basis of the amino acid sequence analysis of a number of glycoproteins, it has been recently proposed that the peptide chain is in a β-turn conformation in the region of glycosylation. This proposal is combined with our spectroscopic data in a model for the glycopeptide conformation in asparagine-linked glycoproteins. The GlcNAc residue is in a $^4C_1$ chair and the amide exocyclic dihedral angles, τ_1 and τ_2, are both 120°. From arguments based on NMR data, both asparagine side chain dihedral angles χ^1 and χ^2 are placed at 60°, and the glycosylated asparagine residue is in the second corner position of a type I β turn.

Although many of the proteins of eucaryotes are glycosylated, the biological functions of the complex oligosaccharide chains of glycoproteins are not generally known. Recent biochemical experiments revealing the extremely complex biosynthetic pathway leading to the N-glycosylation of proteins suggest the biological importance of the carbohydrate moiety (Li et al., 1978; Chen & Lennarz, 1978). A second experimental approach which could provide some insight into glycoprotein function is determination of glycoprotein structure. Although there is considerable progress being made in the covalent chemical structure of glycoproteins, the three-dimensional conformations of glycopeptides have not been extensively discussed.

The amino acid sequence in the region of N-glycosylation of a protein has been found to be of the general type Asn-X-Ser(Thr) with the asparagine side chain amide function N-glycosidically linked to *N*-acetylglucosamine (GlcNAc). Further analysis of the amino acid sequences of glycoproteins by Beeley (1977) suggests that the most likely peptide conformation in the region of N-glycosylation is the β turn. A similar proposal has been made by Aubert et al. (1976).

Certainly, the proposal of a β turn as a general feature for glycosylated regions of glycoproteins seems quite attractive at least for the case of globular proteins, the class which is best represented in the catalogs of amino acid sequence data. One can easily imagine the peptide chain traversing the hydrophobic interior of such a protein, turning at the outer edge to plunge back into the interior, and exposing the glycosylated turn to the aqueous solvent. The highly hydrophilic carbohydrate chain is thus positioned to interact with the aqueous medium as one would expect from simple thermodynamic considerations.

In this paper we will examine certain spectroscopic data on glycopeptide model compounds, chiefly 4-*N*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)asparagine (GlcNAc-Asn). The conformational information deduced from these spectroscopic experiments will be combined with the β-turn conformation for the peptide chain in order to derive a proposed model for the glycopeptide conformation in glycoproteins.

A straightforward conformational analysis of the β pyranose ring suggests a $^4C_1$ chair conformation for the *N*-acetylglucosamine residue in GlcNAc-Asn. This suggestion is supported by X-ray crystallographic data on GlcNAc-Asn (Delbaere, 1974) and by high-field proton NMR spectroscopy in D_2O solution (Dorland et al., 1977). Given the dihedral angles specifying the $^4C_1$ chair, one must then consider three other dihedral angles in order to determine the conformation of the glycopeptide linkage. The asparagine side chain dihedral angles χ^1 and χ^2 represent rotation about the α–β and the β–γ bonds of the amino acid side chain. Assuming planarity for the linkage amide of the asparagine side chain, the third dihedral angle to be defined is that about the glycosyl CN bond.

This exocyclic amide dihedral angle will be called τ_1 and is defined by the four atoms γ-carbonyl carbon, amide nitrogen, anomeric carbon (C1) of the sugar, and C2 of the sugar. A second exocyclic amide dihedral angle (τ_2) describing the orientation of the C2 amide will be defined by C1, C2, the amide 2 nitrogen and amide 2 carbonyl carbon. $\tau = 0$ will be taken as the eclipsed conformation, and positive dihedral angles will correspond to right-handed screw rotations. In a previous paper (Bush & Duben, 1978) these angles were called ϕ_1 and ϕ_2. The present notation will avoid confusion with the standard peptide backbone conformational angles ϕ and ψ.

In a recent theoretical paper on the CD of GlcNAc-Asn, we proposed that assumption of $\tau_1 = \tau_2 = 120°$ gave calculated CD results in agreement with the experimental data. In this paper we present additional experimental data to support that conclusion. Second, we will exploit an analogy between τ_1, τ_2, and the standard peptide backbone conformation angle ϕ_{CN} in proton NMR spectroscopy. In particular, the proton cou-

† From the Department of Chemistry, Illinois Institute of Technology, Chicago, Illinois 60616. *Received June 25, 1979.* This research was supported by National Science Foundation Grant CHE-76-16783.

‡ Present address: Department of Chemistry, South Dakota School of Mines and Technology, Rapid City, SD 57701.

0006-2960/80/0419-0501$01.00/0

pling constant between the amide hydrogen and the pyranose ring hydrogen is related to the dihedral angle by a Karplus-type relation which has been widely used in peptide conformational studies (Ramachandran et al., 1971). We will present NMR data on the amide proton coupling constants which provide further information on the conformation of GlcNAc-Asn in solution.

Experimental Section

Materials. GlcNAc-Asn was purchased from Cyclo Chemical Co. of Los Angeles, CA, and used without further purification. GlcNAc and β-GlcNAc(OAc)$_5$ were purchased from Sigma Chemical Co., St. Louis, MO. α-GlcNAc(OAc)$_5$ was prepared by acetylation of α-GlcNAc in pyridine with acetic anhydride at 4 °C (Horton, 1964). 2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosylamine (I) was prepared by the method of Lis et al. (1974). This compound was acetylated to give 2-acetamido-3,4,6-tri-*O*-acetyl-1-*N*-acetyl-2-deoxy-β-D-glucopyranosylamine (II) (Bolton et al., 1966). Compound II was de-O-acetylated with sodium methoxide in methanol (Bolton et al., 1966) to give 2-acetamido-1-*N*-acetyl-2-deoxy-β-D-glucopyranosylamine (β-1,2-DAG).

CD Spectra. CD spectra were measured in the 170–240-nm region by using a newly constructed vacuum UV CD instrument (Duben & Bush, 1979). Briefly, this instrument uses a short focal length McPherson 218 monochromator and a sealed 200-W deuterium light source. The optics are all under vacuum but the sample chamber at atmospheric pressure is purged with nitrogen. The wavelength range of this instrument is 160–300 nm, the short-wavelength limit being determined by the quartz windows of the light source. Spectra were measured for samples in aqueous solution by using short path length cells. The short-wavelength range is limited to 170 nm by the absorbance of water in the 10-μm path length cells used for this study.

Proton NMR Spectroscopy. Spectra were measured in the Fourier transform mode at 80 MHz on a Varian CFT-20 instrument. Spectra of sugars in ^1H_2O were measured in 0.5 M solution with the instrument operated in the low-field mode and a receiver band width of 500 Hz to minimize receiver overloading by the water resonance. The filter introduced a phase-shift artifact which was corrected by using the internal compensation of the instrument. High-field spectra (360 MHz) of GlcNAc in H_2O were done by Dr. J. Dallas of the Purdue Biochemical Magnetic Resonance Laboratory. Chemical shifts in $CDCl_3$ and in $(CD_3)_2SO$ are reported relative to internal tetramethylsilane. Those in water are referred to internal 3-(trimethylsilyl)propanesulfonate (DSS).

Extinction Coefficients. Since a redetermination of the extinction coefficient of GlcNAc-Asn gave a result differing substantially from that reported by Coduti et al. (1977), the measurement was repeated using different methodology. The concentration of a solution of GlcNAc-Asn was measured by amino acid analysis for aspartic acid after hydrolysis of the sample. From the measured absorption spectrum of the solution a value of 19 000 was determined for the extinction coefficient of GlcNAc-Asn, a value which indicates no hydrochromism. The extinction coefficient of β-1,2-DAG was found to be 15 400.

Results

The CD spectrum of GlcNAc-Asn in water in Figure 1 shows a small positive band in the 210-nm range which is overshadowed by a much stronger band peaking at 198 nm. The appearance of a strong negative CD minimum at 178 nm

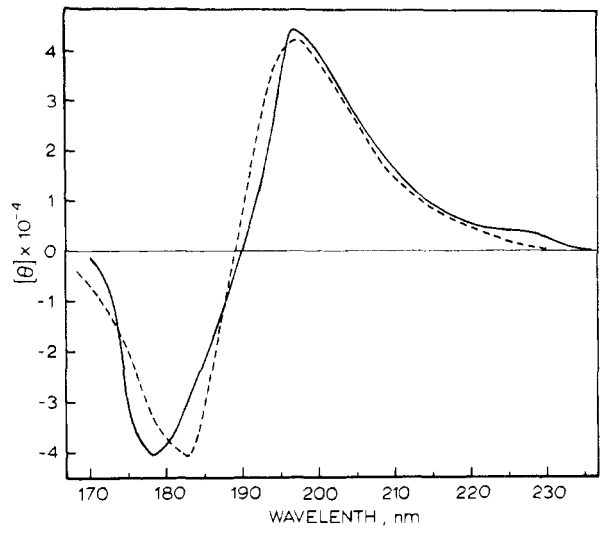

FIGURE 1: Circular dichroism spectra in water of 2-acetamido-1-*N*-(4-L-aspartyl)-2-deoxy-β-D-glucopyranosylamine (—) and of 1-*N*-acetyl-2-acetamido-2-deoxy-β-D-glucopyranosylamine.

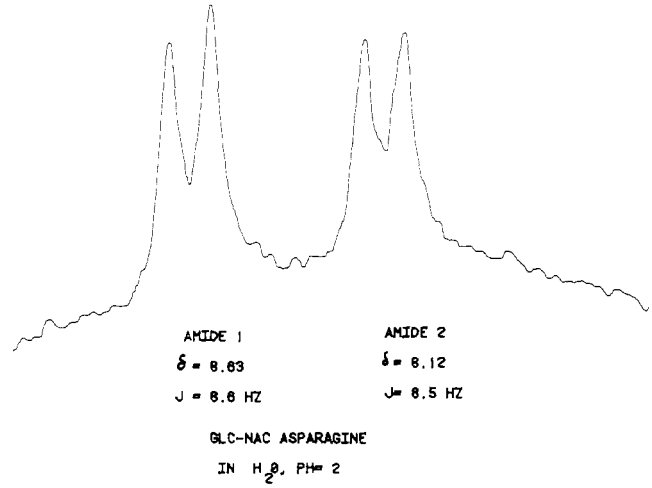

FIGURE 2: Partial NMR spectrum of 2-acetamido-1-*N*-(4-L-aspartyl)-2-deoxy-β-D-glucopyranosylamine in ^1H_2O showing two amide proton resonances.

is in accord with our eariler assignment of these strong CD bands to an exciton pair (Coduti et al., 1977; Bush & Duben, 1978). We had previously identified the exciton coupling with interaction of the vicinal diacyl amino functions of GlcNAc-Asn on the basis of a rather weak dependence of the CD spectrum on pH. The CD spectrum of β-1,2-DAG also shown in Figure 1 is nearly superimposable with that of GlcNAc-Asn, the only significant difference being in the wavelength positions of the minima, perhaps as a result of some chromophoric properties of the amine or carboxylate functions of the latter.

The proton NMR data of Figure 2 on GlcNAc-Asn in water show the two downfield doublets which are assigned to the two amide protons. The assignment is supported by selective homonuclear decoupling experiments. In spectra of this same compound in neutral aqueous solution, the amide 2 proton resonance is slightly broadened while that assigned to amide 1 disappears from the spectrum presumably as a result of rapid exchange with the solvent protons catalyzed by the carboxylate function of the amino acid.

In Table I are summarized more extensive NMR data on acetamido sugars in several solvents. In each case the vicinal coupling constant of the amide proton is large, in agreement with data of Hirano (1972) on various acetamido sugars. It

Table I: Chemical Shifts and Vicinal Coupling Constants of Acylamino Sugars

	shifts and coupling constants		
	peracetylated in $CDCl_3$ δ vs. int Me_4Si	Me_2SO δ vs. int Me_4Si	pH 2, δ vs. int DSS
α-GlcNAc	δ = 5.65	δ = 7.69	δ = 8.01
	J = 9.2 Hz	J = 8.0 Hz	J = 9.3 Hz
β-GlcNAc	δ = 5.79	ND[a]	δ = 8.11
	J = 9.2 Hz	ND	J = 10.0 Hz
β-1,2-diacetamido-glucose			
amide 1	δ = 6.98	δ = 8.53	δ = 8.50
	J = 8.4 Hz	J = 8.6 Hz	J = 8.8 Hz
amide 2	δ = 6.19	δ = 8.01	δ = 8.14
	J = 8.3 Hz	J = 8.3 Hz	J = 8.3 Hz
β-GlcNAc-Asn			
amide 1		δ = 8.43	δ = 8.63
		J = 9.0 Hz	J = 9.0 Hz
amide 2		δ = 7.91	δ = 8.12
		J = 7.3 Hz	J = 8.5 Hz

[a] ND, not determined.

is also noteworthy that the chemical shifts of the amide protons of the α and β anomeric pyranosides differ slightly, the resonance of the β anomer being approximately 0.1 ppm downfield from that for the α anomer.

Discussion

The CD data of Figure 1 confirm our earlier assignment of the large bands in GlcNAc-Asn to an exciton pair resulting from coupling of vicinal amides. The contributions to the CD in this spectral region of the amine and carboxylate chromophores of the amino acid are modest. The exocyclic amide dihedral angles τ_1 and τ_2 must be close to 120° as deduced from CD calculations on models (Bush & Duben, 1978). These values of τ_1 and τ_2 place the amide protons trans to their respective ring protons, a conformation for which large amide proton coupling constants are expected.

Large amide coupling constants have been reported for various acetamido sugars in a wide variety of solvents (Hirano, 1972). Our results summarized in Table I are in agreement with this earlier work and extend the generalization to the technically more difficult but biochemically more relevant solvent water. The fact that large amide proton coupling constants are found in a variety of solvents may be taken to imply that the trans arrangement of the protons is determined primarily by steric exclusion effects rather than by hydrogen bonding or some other polar or solvent-dependent interaction. The small difference in chemical shift of the amide protons of the α and β anomers of GlcNAc presumably arises from slightly different hydrogen bonding of the amide to the solvent. It is possible that this difference in chemical shifts might be exploited in assigning anomeric configuration in oligosaccharide structural studies.

The large coupling constants for both the amide 1 and the amide 2 protons of GlcNAc-Asn suggest dihedral angles τ_1 and τ_2, in agreement with our model based on CD studies. The observation that, in neutral pH solutions, the amide 1 proton resonance is broadened and disappears suggests an interaction between this proton and the carboxylate group of the amino acid. This interaction will be incorporated in our model-building studies below.

In principle, the vicinal α–β proton coupling constants of the asparagine side chain contain information on the side-chain dihedral angle χ^1. Such data have been obtained from high-field proton NMR spectroscopy by Dorland et al. (1977). These coupling constants may be used to compute the populations of the three rotomers expected for χ^1 by using the method of Feeney (1976). Unfortunately, since the two β protons have not been explicitly assigned, there are two possible solutions. If we assign the low-field proton (δ = 2.93; J = 4.00 Hz) as H_A in Feeney's notation, we conclude that the conformation should contain essentially an equal mixture of χ^1 = 60° and χ^1 = −60°. If H_A is taken as the proton at higher field (δ = 2.87; J = 6.90 Hz), the most probable value of χ^1 is 60° with χ^1 = 180° also contributing. Either of the two assignments leads to the conclusion that the model compound GlcNAc-Asn has conformational flexibility about the side-chain dihedral angles but that χ^1 = 60° is a probable value. Model-building studies show that such a side-chain conformation can be stabilized in a glycopeptide by hydrogen bonding between the side-chain amide and the peptide backbone in a type I β turn. Of two likely types of hydrogen bonding, one involving bonding of the side-chain amide oxygen to the amide proton of the peptide chain in a six-membered ring has been discussed by Kopple & Go (1977). A second possible hydrogen-bonding interaction, that between the side-chain amide proton and the peptide oxygen, is suggested by our observation that the carboxylate oxygen of GlcNAc-Asn catalyzes the exchange of the amide 1 proton with the solvent. It is the second type of hydrogen bond that we have incorporated in the model.

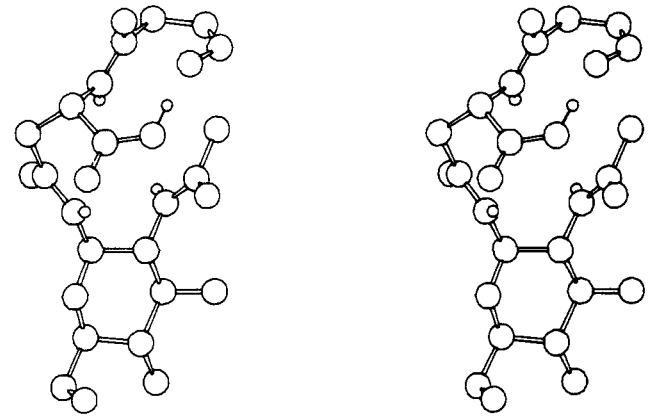

FIGURE 3: Stereo view of the conformational model of asparagine-linked glycopeptides.

The proposed model illustrated in Figure 3 features the pyranoside in a $^4C_1$ chair. The amide dihedral angles τ_1 and τ_2 are taken as 120°, in agreement with both CD and NMR data. The asparagine side chain dihedral angle χ^1 is set at 60° on the basis of the α–β proton coupling constants as discussed above. χ^2 is set at +60° in order to provide interaction between the amide 1 proton and the carboxylate oxygen of the amino acid chain to form a seven-membered ring closed by a hydrogen bond. Although the peptide conformation in the region of N-glycosylation is likely to be a β turn, Beeley (1977) points out that the exact identities of the turn residues are not clearly defined by the method of amino acid sequence analysis. Nevertheless, it is most commonly the second corner residue, however, which is glycosylated (Beeley, 1977; Aubert, 1976). Of the two common types of β-turn conformation, sequences with asparagine in the second corner position generally form the type I turn (Kopple & Go, 1977; Chou & Fasman, 1977). Just enough amino acid residues are included in the model of Figure 3 to show the hydrogen-bonding interaction of the ten-membered ring of the type I β turn. Although the model proposed in Figure 3 is somewhat spec-

ulative, there is at least some physical evidence for each of its features. These features are generally susceptible to test by further spectroscopic studies. Experiments and conformational calculations are now in progress to test these various features.

Acknowledgments

We thank Brian Rogers for the synthesis of 1,2-DAG. Thanks are also due to Dr. K. D. Kopple for help with the NMR measurements and to Dr. J. Dallas for the high-field spectroscopy.

References

Aubert, J. M., Biserte, G., & Loucheux-Lefebvre, M. H. (1976) *Arch. Biochem. Biophys. 175*, 410.
Beeley, J. G. (1977) *Biochem. Biophys. Res. Commun. 76*, 1051.
Bolton, C. H., Hough, L., & Khan, M. Y. (1966) *Biochem. J. 101*, 184.
Bush, C. A., & Duben, A. J. (1978) *J. Am. Chem. Soc. 100*, 4987.
Chen, W. W., & Lennarz, W. J. (1978) *J. Biol. Chem. 253*, 5775.
Chou, P., & Fasman, G. D. (1977) *J. Mol. Biol. 115*, 135.
Coduti, P. L., Gordon, E. C., & Bush, C. A. (1977) *Anal. Biochem. 78*, 9.
Delbaere, L. T. J. (1974) *Biochem. J. 143*, 197.
Dorland, L., Schut, B. L., Vliegenthart, J. F. G., Strecker, G., Fournet, B., Spik, G., & Montreuil, J. (1977) *Eur. J. Biochem. 73*, 93.
Duben, A. J., & Bush, C. A. (1979) *Anal. Chem.* (in press).
Feeney, J. (1976) *J. Magn. Reson. 21*, 473.
Hirano, S. (1972) *Agric. Biol. Chem. 36*, 1071.
Horton, D. (1964) *J. Org. Chem. 29*, 1776.
Kopple, K. D., & Go, A. (1977) *J. Am. Chem. Soc. 99*, 7698.
Li, E., Tabas, I., & Kornfeld, S. (1978) *J. Biol. Chem. 253*, 7762.
Lis, H., Lotan, R., & Sharon, N. (1974) *Methods Enzymol. 34*, 341.
Ramachandran, G. N., Chandrasekaran, R., & Kopple, K. D. (1971) *Biopolymers 10*, 2113.

Stabilization of Nucleic Acid Secondary Structure by Cationic Metal Complexes†

Richard L. Karpel,‡ Arthur H. Bertelsen, and Jacques R. Fresco*

ABSTRACT: Complexes of inert (slowly exchanging) multipositive transition metal cations and neutral ligands very effectively stabilize both RNA and DNA double helices against thermal denaturation. This indicates that these cations prefer binding sites on double helices over those on single strands. The preferential affinity of the tripositive complexes for the double helices poly(I·C) and poly(U·U) is so great that it results at low [cation]/[polynucleotide phosphate] ratios in biphasic melting of these helices. The interaction between the complex cations and helical nucleic acids has a strong electrostatic component since T_m enhancement is diminished with increasing sodium ion concentration. In addition, the trivalent complexes, $Co(NH_3)_6^{3+}$, $Co(en)_3^{3+}$, and $Pt(en_2)(en\text{-}H)^{3+}$, show greater enhancement of T_m than the divalent ones, $Ir(NH_3)_5Cl^{2+}$ and $Pt(NH_3)_4^{2+}$. These tripositive complexes are similarly effective in stabilizing poly(I·C), suggesting that the helical binding sites for these cations are less discriminating than the sites to which they bind in stabilizing the tertiary structure of $tRNA_3^{Leu}$ [Karpel, R. L., Miller, N. S., Lesk, A. M., & Fresco, J. R. (1975) *J. Mol. Biol. 97*, 519-532]. On the other hand, $Ir(NH_3)_5Cl^{2+}$ was more effective in stabilizing poly(I·C) than was $Pt(NH_3)_4^{2+}$. The two groups of complexes had a lesser effect on the thermal stability of DNA helices. A quantitative treatment of the dependence of T_m on complex cation concentration is used to obtain association constants and the helical binding site size.

Nucleic acids, by virtue of their polyanionic nature, require the presence of cations to stabilize double-helical secondary structure. Cations can reduce the electrostatic energy of these molecules through two different mechanisms, by Debye-Huckel shielding of the phosphates and by site binding to them [cf. Felsenfeld & Miles (1967)]. The requirement for cations is well illustrated by the variation of the thermal denaturation of DNA with the concentration of shielding monovalent alkali metal ions such as Na^+ and K^+, which shows a linear dependence of the melting temperature, T_m (the temperature at the midpoint of the cooperative helix-coil transition), on log cation concentration (Schildkraut & Lifson, 1965). Obviously, diminution of the electrostatic repulsion of interstrand phosphates by cations increases the stability of the helix. Indeed, divalent alkaline earth metal ions such as Mg^{2+} or Ca^{2+}, which site bind (only) to the phosphate moieties, increase the T_m of nucleic acid helices even more dramatically (Dove & Davidson, 1962).

All metal ions do not, however, affect nucleic acids in this manner. Several transition metal ions with an affinity for the phosphates exhibit complex dependencies of T_m on cation concentration because of the additional tendency to coordinate to purines and, to a lesser extent, to pyrimidines (Izatt et al., 1971; Eichorn, 1973). Some of these cations, e.g., Cu^{2+}, even unwind nucleic acid helices because they cannot be accommodated in the helix when they bind tightly to the bases (Eichhorn & Shin, 1968; Eichhorn & Tarien, 1967), while others, such as Zn^{2+} and La^{3+}, also catalyze the scission of

† From the Department of Biochemical Sciences, Frick Laboratory, Princeton University, Princeton, New Jersey 08544. *Received July 27, 1978; revised manuscript received June 1, 1979.* This work was supported by grants from the National Institutes of Health (GM-07654) and the National Science Foundation (GB-35595). This is paper 16 in a series entitled "Polynucleotides" of which the previous paper is Topal & Fresco (1974).

‡ R.L.K. was a U.S. Public Health Service Postdoctoral fellow (1974-1976). Present address: Department of Chemistry, University of Maryland, Catonsville, MD 21228.

0006-2960/80/0419-0504$01.00/0