TAUTOMERISM OF NUCLEIC ACID BASES AND PMR STUDIES OF HISTONES

I. PROTON MAGNETIC RESONANCE STUDIES OF TAUTOMERISM OF NUCLEIC ACID BASES

II. PROTON MAGNETIC RESONANCE STUDIES OF HISTONES AND A DNA-HISTONE COMPLEX

III. NUCLEAR MAGNETIC RESONANCE STUDIES OF RIBONUCLEASE A- INHIBITOR INTERACTIONS AND NUCLEOTIDE CONFORMATIONS

Thesis by

CHI-YU GREGORY LEE

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1972

(Submitted January 18, 1972)

ACKNOWLEDGEMENTS

First, I would like to give many thanks to my research adviser, Professor Sunney I. Chan, without whose guidance, this thesis would not have reached completion. His deep understanding and interest in the problems encountered throughout the research have encouraged me to explore their nature with extensive efforts. His motto, "Always do the experiment which you think should not work," has made my research fruitful and successful. I am also indebted to Professor Norman Davidson for his enthusiastic discussions about the results of this research. I should also like to express my sincere gratitude to my former groupmate, Dr. J. H. Prestegard. Without his help at the initial stage of my graduate research, I could not have solved this problem successfully. Another groupmate , Dr. G.P. Kreishman's discussion about the biological implications of this problem is also highly appreciated. I must also acknowledge the support from the other members of the Chan research group whose help was so valuable during my graduate study at Caltech. Finally, I wish to express my appreciation to the excellent thesis typing of Mrs. Pierce and to Caltech for the financial support of my graduate study.

Dedicated to my parents

Mr. A. Y. Lee and Mrs. S. Kou Lee

ABSTRACT

PART I. Proton Magnetic Resonance Studies of Tautomerism of Nucleic Acid Bases

1. Cytosine. It has long been noted that the H_5 proton of cytosine and its related derivatives exhibits unusually broad resonances in the pmr spectrum under certain conditions of temperature and solution pD. We have examined this phenomenon as a function of temperature, concentration, and solution pD as well as the external magnetic field, and have shown that the linebroadening arises from chemical exchange between the amino and imino tautomers of the cytosine base. The observation of sharp H_5 resonances in cytosine derivatives existing in only the amino or imino tautomeric structure supports this interpretation. This tautomeric exchange was found to be catalyzed by OD^- as well as the solvent molecules (D_2O). Analysis of the temperature and pD dependences of the H_5 linewidth for 5'-CMP led to the following rate law at 30°C:

$$-\left(\frac{d\,A}{dt}\right)_f = 1.2 \times 10^{10}\ [OD^-]\ [A] + 1.2\ [D_2O]\ [A], (Msec^{-1})$$

where [A] represents the major tautomeric species of 5'-CMP. Activation energies of 3.5 kcal/mole and 7.8 kcal/mole were obtained for the OD^- and D_2O catalyzed steps respectively. The imino tautomer was estimated to be present to the extent of 15 ± 3% at room temperature in neutral aqueous solution.

2. Guanine. It has long been noted that the H_8 proton of guanosine and its related derivatives exhibits unusually broad resonances in the pmr spectra under certain conditions of temperature and pD. We have examined this phenomenon as a function of temperature, concentration, and solution pD, as well as the external magnetic field, and have shown that the linebroadening arises from chemical exchange between the lactam and lactim tautomers of the guanine base. The observation of sharp H_8 resonances in guanine derivatives where the guanine base is frozen in only the lactam and lactim tautomeric structure supports this interpretation. This tautomeric exchange was found to be catalyzed by OD^- and the solvent (D_2O). Analysis of the temperature and pD dependence of the H_8 linewidth for 2'-GMP led to the following rate law at 30°C:

$$-\left(\frac{d\,A}{dt}\right)_f = 1.7 \times 10^9 \left[OD^-\right]\left[A\right] + 0.7\left[D_2O\right]\left[A\right](M\ sec^{-1}),$$

where A represents the major tautomeric species of 2'-GMP. Activation energies of 7 kcal/mole and 13 kcal/mole were deduced for the OD^- and D_2O catalyzed steps respectively. The minor lactim tautomer was estimated to be present to the extent of 16 ± 3% at room temperature in neutral aqueous solution.

In most guanosine derivatives, with the notable

exception of 2'-GMP, the H_8 linewidth was also found to be strongly concentration dependent over the pD range 3 to 6. This concentration dependence has been interpreted in terms of the effect of intermolecular association on the lactam-lactim tautomeric equilibrium. Analysis of the data in terms of a rapid equilibrium involving the monomer, a hydrogen-bonded tetramer and stacked aggregates of this tetramer yielded a tetramer formation constant of $(2.5 \pm 0.5) \times 10^7\ M^{-3}$ and a tetramer stacking or association constant of $40 \pm 10 M^{-1}$ in the case of 5'-GMP.

PART II. Proton Magnetic Resonance Studies of Histones and A DNA-Hostone Complex.

1. Solution Properties of Nucleoproteins - Histone I and Histone IIbI. The 220 MHz high resolution proton magnetic resonance (pmr) spectra of histone I and histone IIbl have been examined as a function of histone concentration, salt concentration as well as solution pD. The concentration studies of both histone I and histone IIbl revealed exclusively monomer-dimer equilibria in the presence or the absence of salt. The increase in the dimer formation of the proteins is attributed to the counter ion binding upon the addition of salt which results in shielding of the long range electrostatic repulsion and facilitating the hydrophobic interactions between the

molecules. Changes in moles of the counter ion $\Delta\bar{\nu}_x$ and water $\Delta\bar{\nu}_w$ bound to each mole of histones upon the complex formation were determined experimentally. Temperature studies of histone I and histone IIb1 indicate that in the presence of salt, the entropy change for the dimer formation of these proteins, ΔS becomes more positive than that in the absence of salt.

2. Proton Magnetic Resonance Studies of the DNA-Histone I Complex. The DNA-histone I complex prepared by a continuous salt gradient dialysis was studied by high resolution proton magnetic resonance spectroscopy at various temperatures. It was concluded that about 15 to 20% of the side chain protons of the amino acid residues of the complexed histone I can be observed in the pmr spectra, and histone I in the complex is not removed from DNA when the latter is thermally denatured. The spectral behaviors of histone I when mixed directly with DNA or poly U were examined at various temperatures for the purpose of comparison. Possible structural models of this DNA-histone I reconstituted complex are presented on the basis of our experimental observations.

PART III. Nuclear Magnetic Resonance Studies of Ribonuclease A Inhibitor Interactions and Nucleotide Conformations.

1. A ^{31}P NMR Study of Interactions of Ribonuclease A and its Inhibitor . The association between ribonuclease A and its product inhibitor uridine-3'-monophosphate has been studied by ^{31}P nmr spectroscopy. Analysis of chemical shifts and the linewidths of the ^{31}P resonance of the inhibitor over a wide range of enzyme concentrations yielded the apparent dissociation rate constant 3200 ± 300 sec^{-1} at 32°C for the enzyme-inhibitor complex and an activation energy of 3.5 ± 1.0 kcal/mole for this process. The chemical shift data suggest that the phosphate group of this inhibitor is simultaneously bound to two positively charged amino acid residues at the active site of the enzyme.

2. A Proton Magnetic Resonance Study of the Effects of pH on Nucleosides and Nucleotides in Aqueous Solution. Experimentally it was observed that $J_{H_1'-H_2'}$ of some guanine nucleosides and nucleotides varies with pD of the solution. This phenomenon was interpreted on the basis of the ring-base interactions due to the protonation of the guanine base at low pD's. From the results of the

conformational analysis of these guanine derivatives, it was concluded that guanosine and 3'-GMP and 5'-GMP exist predominantly in syn conformation when the guanine base is protonated at low pD's, while the anti conformation is still preferable at high pD's when the base is neutral.

TABLE OF CONTENTS

PART I

PROTON MAGNETIC RESONANCE STUDIES OF TAUTOMERISM OF NUCLEIC ACID BASES

I. INTRODUCTION.

1. Role of DNA in the Fundamental Living Processes.

The most striking attribute of a living system is its ability to transmit the hereditary properties from one generation to another. The regularity of the heredity led to the discovery of its basic rule by Mandel in 1865 who first proposed that various traits of an individual are controlled by pairs of factors which are derived from male and female parents respectively and initiated the field of genetics.[1] With the passage of time, genetics has now progressed to the stage where its concepts and phenomena can be described on the molecular level.[2] The major breakthrough is the discovery of the double helical structure of deoxyribonucleic acid (DNA), the basic hereditary molecule,[3-5] by Watson and Crick in 1953. Based on this particular structure model, it is possible to explain how this molecule duplicates itself accurately and how the genetic information carried by this molecule can be exactly transmitted into an amino acid sequence of a given protein.[3-5]

According to the Watson and Crick model, DNA molecule is a double helix with two sugar-phosphate chains held together by a sequence of hydrogen-bonded pairs of nucleotide bases. In a typical DNA molecule, there are only four distinguishable component bases, namely the

pyrimidine bases; thymine and cytosine, and the purine bases; adenine and guanine.

The pairing of the bases between the two sugar-phosphate chains in DNA is very specific, i.e., thymine pairing only with adenine and cytosine pairing with guanine only. Because of this unique base pairing, the two strands of a double helical DNA molecule are exactly complementary to each other.

During the cell replication, the two strands of DNA are separated. Each of the two strands of DNA acts as a template for the enzymatic synthesis of a new DNA molecule using monomeric 2'-deoxyribonucleoside-triphosphates as primers for the synthesis of a new strand of DNA. This newly synthesized DNA strand should be complementary to the old strand according to the unique base pairing restriction imposed between the pyrimidine and purine bases. Thus each half of the helix is forming its own couple and as a result two identical DNA molecules are accurately formed after the replication.(2)

Compared to DNA replication, the mechanisms of how DNA controls the amino acid in protein are much more sophisticated. Since DNA itself does not act as a direct template that orders amino acid sequence, the genetic information carried and determined by sets of particular base sequence in this molecule must be transferred to another class of molecules which then serve as the protein

templates.[6] These intermediate templates are molecules of ribonucleic acid (RNA). The transcription of RNA from the DNA template shows many similarities to the DNA replication. The enzyme RNA polymerase links together the monomeric precursors which are ribonucleoside-triphosphates (ATP, GTP, CTP and UTP) to form a single stranded RNA complementary to that segment of the two DNA strands which corresponds to a given gene. There are important differences between DNA and RNA. RNA chains are much shorter--contain 600 bases, if it is to direct the synthesis of a protein of molecular weight of 25000; one of the nucleic acid bases, thymine, is replaced by uracil and the sugar backbone is slightly different--ribose ring instead of deoxyribose. Through this RNA intermediary, the message is translated into specific sequences of 20 different amino acids in the protein chain. The messenger RNA which has been synthesized complementary to the appropriate base sequence in the DNA literally carries the DNA message to the ribosome, when it sets and starts the step-by-step synthesis of the polypeptide chain. There are in the cytoplasm of cells small transfer RNA's (80 base units for each and 20 of these), each of which is capable of recognizing one of the 20 different amino acids and transferring it to the site where it is to be incorporated into a growing polypeptide chain. It is now known that each successive triplet of bases on the messenger RNA (codon) recognizes a complementary

triplet of bases on the tRNA (anticodon), and each codon specifies a particular amino acid (1:1 correspondence between codon and amino acid).

Thus, at each step in the process of protein synthesis, a crucial role is played by the various RNA's. Messenger RNA transcribes the genetic message for each protein from its storage site in DNA. Another kind of RNA, ribosomal RNA which forms part of the structure of ribosomes, acts as a jig for holding the messenger RNA in place while the message is transcribed into the synthesis of polypeptide chain via tRNA's, each of which is capable of recognizing one of the 20 amino acids. As a result, the synthesized protein corresponds exactly to that described by each gene of the DNA molecules.

From the aspect of molecular genetics, the nucleic acid bases therefore are the basic genetic materials and really play an important role in the fundamental living processes. A better understanding of the chemical structures and physical properties of these nucleic acid bases *in vitro* will lead to a better interpretation of the genetics on the molecular and submolecular level.

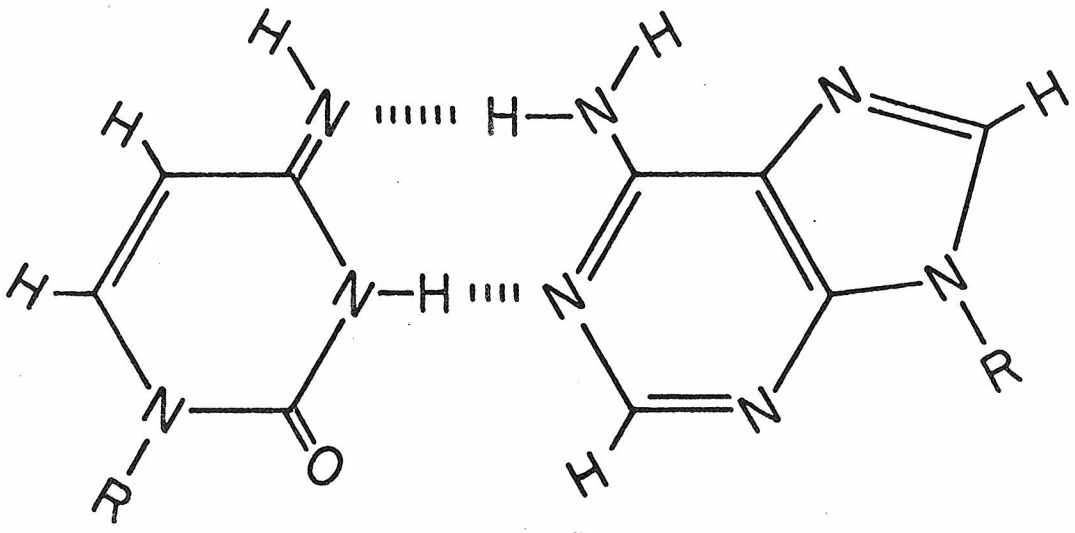
2. Biological Implication of the Tautomerism of Nucleic Acid Bases.

It is well known that the accuracy of the transfer of genetic information during DNA replication and RNA

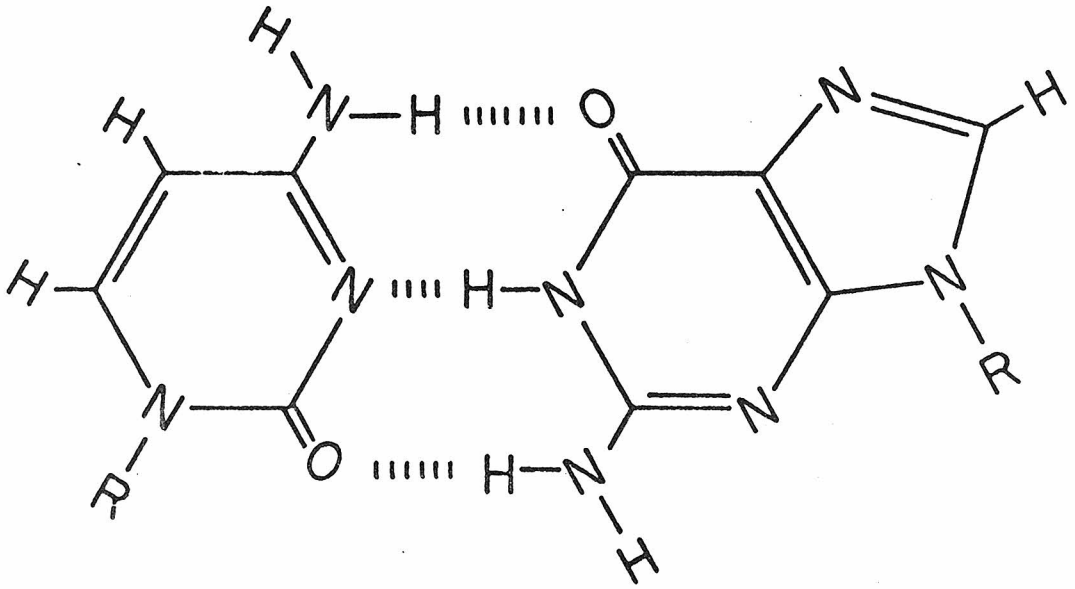
transcription relies on the unique base pairing of the complementary nucleic acid bases,[3,5] i.e., adenine-thymine (or uracil) and guanine-cytosine base pairing. However, these bases may occasionally exist in other rare tautomeric forms in addition to the principal tautomeric structures, and it has been postulated that the presence of these minor base tautomers at the time of DNA replication or RNA transcription would be an important source of imperfect base pairing which can lead to genetic mutation.[8] For example, the lactim form of the guanine (G) base has an electronic structure which is superimposable with that of the normal adenine (A) base,and which may be misread and paired with a thymine (T) base to form an abnormal G-T base pair instead of the normal Watson-Crick G-C base pair (Fig. 2). Similarly, the cytosine (C) base may also exist in either its normal amino or abnormal imino form. It is well known that the normal amino form pairs with guanine base to form stable Watson-Crick G-C base pair. The imino tautomer, by contrast, has the appropriate electronic structure for pairing with the adenine base to form the abnormal A-C base pair (Fig. 1).

The tautomerism of nucleic acid bases can also lead to alternate base pairing schemes other than the normal Watson-Crick schemes for codon-anticodon interactions. The codon-anticodon interaction during the protein synthesis is known to be degenerate.[2,9,10] This degeneracy shows

Figure 1. Abnormal A-C pair and normal G-C pair.

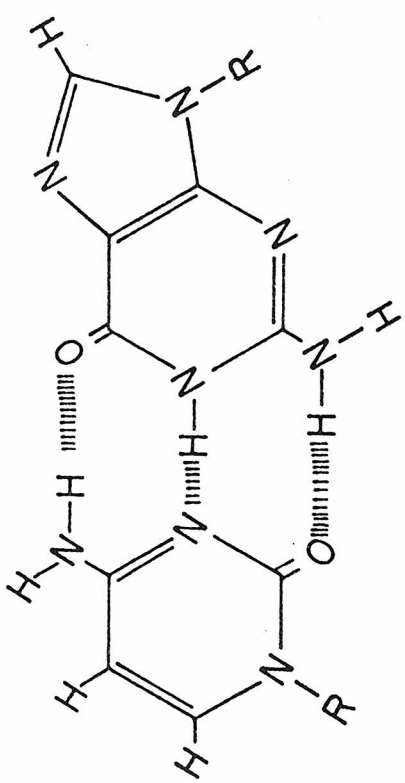


C — A

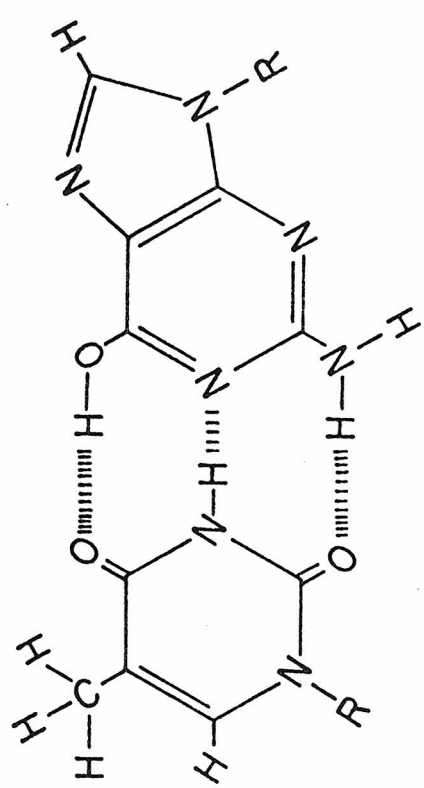


C — G

Figure 2. Normal G-C pair and abnormal G-T pair.



C — G

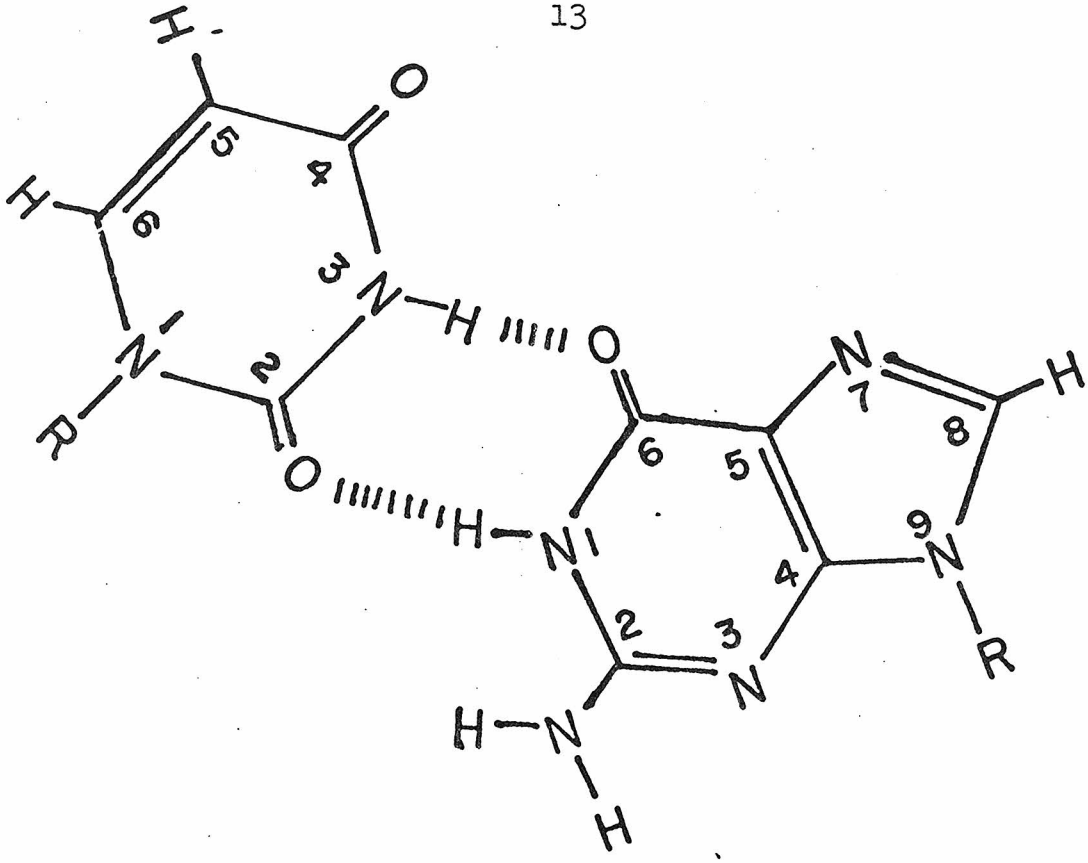


T — G

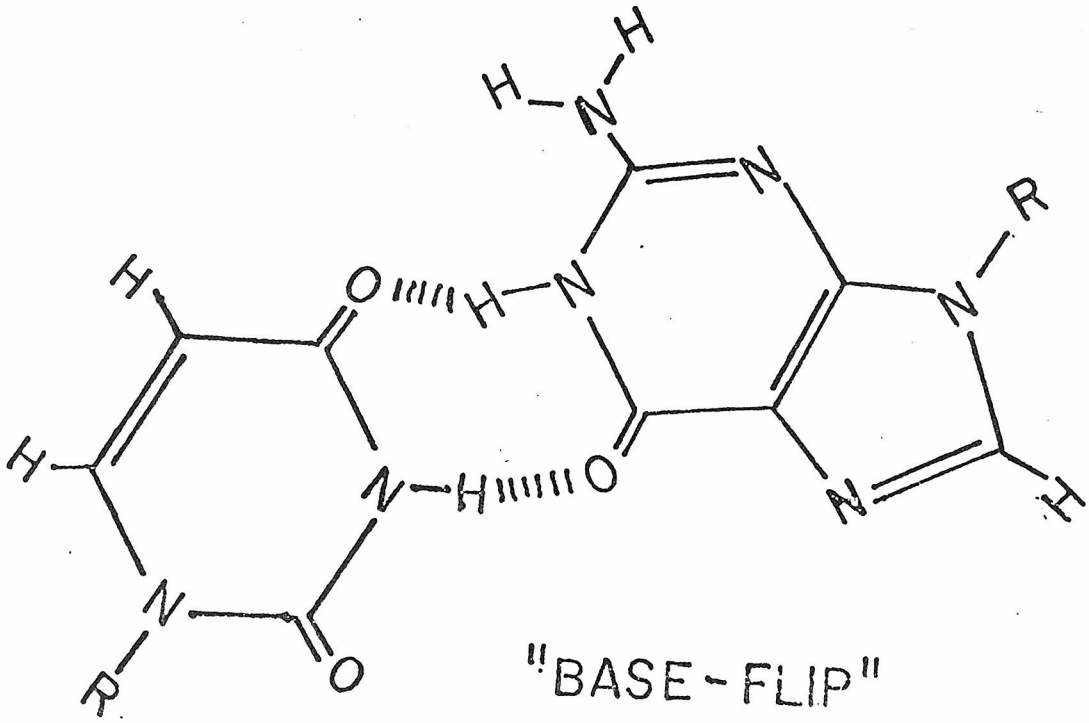
marked inherent regularity.[11] The first two letter bases are very specific and should be complementary. The third letter base, however, is equivalent with U = C for 16 of the 16 possible cases of the first two letters and is also equivalent for A = G for 14 of the 16 possible cases. It is interesting to note that the degeneracy of the code lies in the codons and not in the anticodons, i.e., a single tRNA can interact with several codons in mRNA, for example, in doublet XYU and XYC or XYA and XYG.[12-14]

The mechanism of the interaction of a single tRNA molecule with several degenerate codons is still unknown. Several theories, however, have been proposed to explain this phenomenon, i.e., (1) "wobble" hypothesis,[15] (2) "base flip" hypothesis,[16] (3) "tautomeric doublet" hypothesis.[17] "Wobble" hypothesis proposes that the third base of the codon is allowed to assume several different orientations enabling it to exhibit several hydrogen bonding schemes. The "base flip" hypothesis states that the third base can be either in a syn or anti conformation which allows this base to have two distinct hydrogen bonding schemes. The "tautomeric doublet" hypothesis states that the third base of codon is more exposed to the solvent and can exist in two exchangeable tautomeric forms which allow this base to form two different base pairings(Fig. 3). In view of the last proposal, the possibility that the tautomerism of these

Figure 3. "Wobble" G-U pair and "base-flip" G-U pair.



"WOBBLE"



"BASE-FLIP"

bases may be involved in the degeneracy of genetic code should also be considered.[18-20]

In fact, the importance of these alternate base pairing schemes has been demonstrated by a model study of the interaction of poly U with ApApA and ApApG in aqueous solution.[21] In the study of the interaction between poly U and ApApA, Kreishman and Chan have shown that although below room temperature a rigid triple-stranded complex involving a stoichiometry of one adenine to two uracil bases is more favorable, a complex of 1:1 nature with Watson-Crick base pairing scheme is formed between ApApA and poly U at high temperatures. From the study of the interaction between poly U and ApApG, they have indicated that the complex is 1:1 nature which exhibited the Hoogsteen bonding scheme at low temperatures and the guanine base of ApApG appears to be in the lactim tautomeric form in the complex. In view of this observation, the "tautomeric doublet" hypothesis might be a reasonable explanation of the degeneracy of codon-anticodon interactions between tRNA and mRNA during protein synthesis.

3. Historical Background of the Research--Previous Work .

Suffice it to say, numerous attempts have been made over the past decade and a half, both experimental and theoretical, to ascertain the existence of the tautomerism in nucleic acid bases and to determine the energetics

involved.

3.1. Experimental Work .

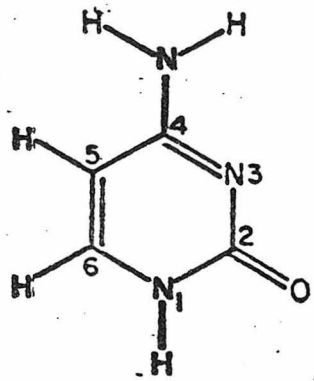
Experimentally, the question of tautomerism of nucleic acid bases has been investigated by various experimental techniques including x-ray crystallography, uv, ir, Raman and nmr spectroscopy, and thermodynamic pK_a measurements.

Among the various methods employed, x-ray crystallography is the most direct and reliable method in determining the structures of molecules in the solid state.[22] The results of these studies on the nucleic acid bases indicated that the amino and/or lactam tautomeric forms of these bases exist predominantly in the crystalline state. Although the solid state structures may not be exactly the same as those in solution, the amino and/or lactam tautomeric forms of nucleic acid bases are generally accepted as the predominant species in solution.

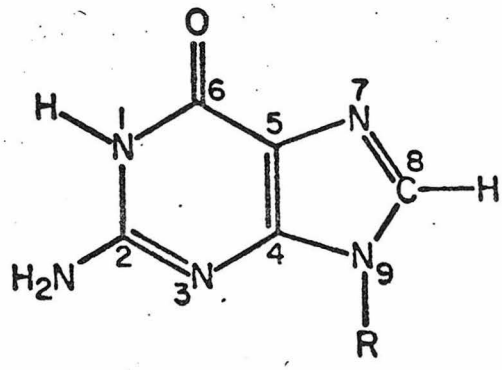
UV studies have been made by the comparison of the extinction coefficients of the uv absorption spectra at various wavelengths among the base and its alkylated analogs which are frozen in either tautomeric form. If the alkyl substitution on the base does not change spectra and the uv extinction coefficient of the base, the observed uv extinction coefficients of the base at various wavelengths can be assumed to be the superposition of those of the alkylated bases frozen in either tautomeric form.

In the study of cytosine, e.g., Kenner et al.[23] compared the uv spectra of the cytosine base with its alkylated analogs 1,4-N, N trimethyl cytosine and 1,3 dimethyl cytosine in alcoholic solution and concluded that the amino to imino ratio of the base is $10^{4.7}$. Similar uv studies by other workers showed that uracil, guanine and adenine exist predominantly either in lactam or amino tautomeric form in aqueous solution.[24-30] Katritzky et al. concluded from the quantitative uv studies that the ratios between lactam and lactim tautomeric forms of uracil and 5-bromouracil are $10^{3.3}$ and $10^{1.7}$ respectively.[24,25] These results are also consistent with those of the pK_a measurements among the base and its methylated analogs. A recent uv study of Wolfenden has shown that the amino-imino ratio of the adenine base is $10^{4.0}$, but that the lactam-lactim ratio of the inosine is about 1.0[27] However the result of this inosine study is not in agreement with that of a Raman study recently reported by Medeiros and Thomas.[31] From the comparison of the Raman spectra of inosine and its lactam as well as lactim methylated analogs in the region 200 to 2000cm^{-1}, Medeiros and Thomas concluded that inosine exists predominantly in the lactam form rather than the lactim form in aqueous solution.

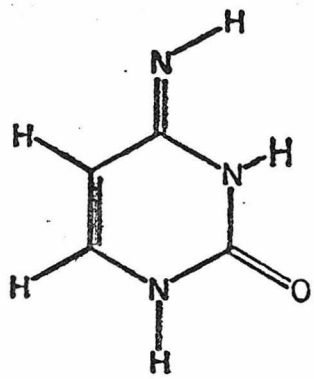
Recent uv studies on the cytosine base by Hélène et al.[32] as well as Morita et al.[33] have indicated



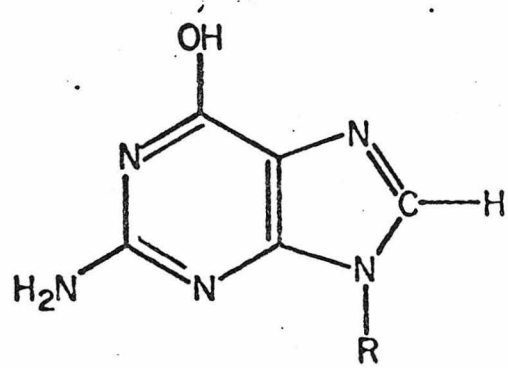
Cytosine (Amino)



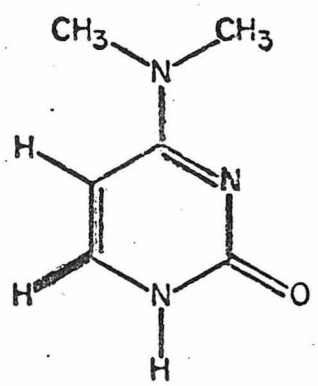
Guanosine (Lactam)



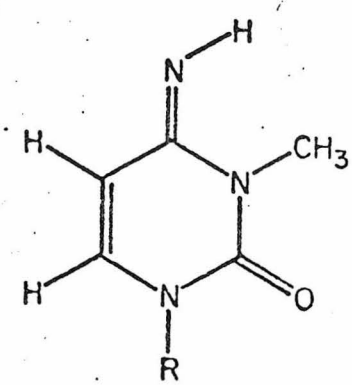
Cytosine (Imino)



Guanosine (Lactim)



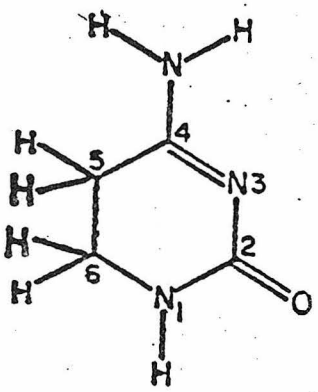
4-N, N dimethyl Cytosine



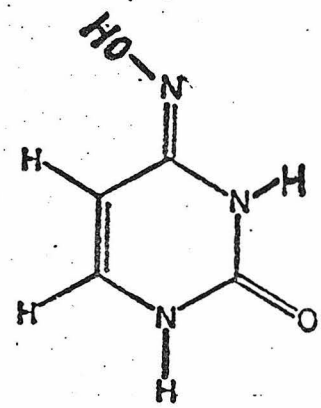
3-methyl,
4-imino Cytidine
7.5 < PD < 9.5

that this base and its analogs (such as isocytosine) undergo a large change in their electronic absorption spectra on going from aqueous to nonaqueous solution and concluded that there are two tautomeric species involved, one predominantly in aqueous solution and another in nonaqueous solution. From the temperature dependence of the uv spectra of these compounds, the difference in free energy between these two species has been estimated to be 2 kcal/mole which corresponds to a tautomeric equilibrium constant of $K_T \approx 30$. Morita and Nagakura have recently suggested the imino form of the cytosine base as a tautomer corresponding to the second species.[33] This conclusion has been opposed by a similar study of Johnson et al.,[34] since the latter argued that the difference in uv absorption spectra in both aqueous and nonaqueous solution may be due to the difference in the solvation of the base and that the second species indicated by Morita et al. could be the hydrogen-bonded solvation complex of the cytosine base. On the basis of the uv work of Johnson et al., it thus appears that the normal amino form of the cytosine base predominates in neutral aqueous solution.

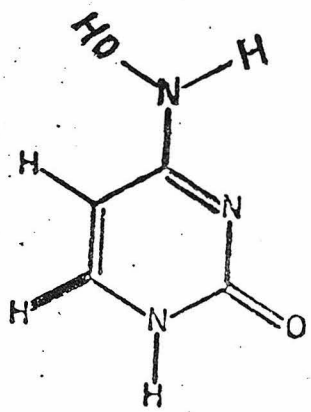
The tautomerism of some cytosine derivatives has also been investigated in detail.[35,36] In aqueous solution, 4-N, hydroxy-cytosine and its 1-N derivatives exist predominantly in oximino form with a tautomeric equilibrium constant (oxamino-oximino ratio) of 0.1, which



5,6 dihydrocytosine



4-N, oximino cytosine



4-N, oxamino cytosine

was determined by uv, ir studies as well as pK_a measurements. A similar study on 1-methyl 5,6 dihydrocytosine has indicated that this compound exists predominantly in the amino form with amino-imino ratio of 25 in contrast with the value of $10^{4.7}$ for cytosine. However, a large shift toward imino form was observed in chloroform solution, where the equilibrium amino-imino ratio becomes 0.1.

Although uv spectroscopy is a highly sensitive tool and is amenable to work in aqueous solution, the similarity of the spectra and overlap of the component bands of many nucleic acid bases as well as their alkylated analogs make its application to the quantitative structural determination unreliable.

The determination of the tautomeric equilibrium constants of nucleic acid bases from thermodynamic K_a measurements has also been reported.[23-25,35-37] If the alkyl substitution does not affect the protonation constant of the base, and if the common cation exists upon the protonation of the base as well as its alkylated analogs, then the tautomeric equilibrium constant K_T would be roughly equivalent to the ratio between the K_a's of the alkylated base analogs, i.e., $K_T \approx K_a^{I}/K_a^{II}$, where K_a^{I} and K_a^{II} are the protonation constants of the akylated base analogs frozen in tautomeric form I and II respectively. However the results of this study are very inaccurate, in view of the possible K_a shift upon the alkylation of the

bases and multiplicity of the ionizable group in these bases.

Infrared spectroscopy has long been applied to the structural determination of nucleic acid bases[38-43] In these studies specific vibrational absorption bands of the functional groups such as C = O C - OH, C - NH_2, C = NH were carefully examined among various bases and their alkylated analogs in the wavelength region from 2 μ to 25 μ in D_2O solution (to avoid the strong interference from water). Extensive studies on the tautomeric structural determination of cytosine, inosine as well as other nucleic acid bases have been reported by T. Miles. In the case of cytosine base, the ir spectra of cytidine was compared with those of 3-methyl as well as 4-N,N dimethyl cytidine derivatives in the wavelength region 1550 cm^{-1} to 1750 cm^{-1}. The amino group has strong absorption band at 1651 cm^{-1} and the imino group has strong absorption band at 1671 cm^{-1} as well as moderate absorption band at 1579 cm^{-1}. On the basis of the ir data, cytidine was shown to exist predominantly in the amino form. Similarly, inosine was compared with 1-methyl as well as 6-methoxy inosine in the same ir absorption region. The absence of the enol group absorption in inosine suggests that this base exists predominantly in the lactam form in solution. The studies on the other bases also showed that they exist predominantly either in amino form or lactam form.

Although ir method is adequate for the qualitative distinction among various tautomeric forms of the bases, the quantitative study is limited by the serious overlap of the absorption bands and by unknown alkyl substituent effects which cause shifts in the component absorption bands as well as the changes in the extinction coefficient of the ir absorption bands.

The structure of nucleic acid bases has also been studied by nmr spectroscopy both in aqueous and nonaqueous solvents.[44-47] In nonaqueous solvent such as dimethyl sulfoxide and chloroform, the resonances arising from protons on NH_2, NH or OH group on the nucleic acid bases can be observed. However apparently only the resonances of major tautomers were detected in these spectra. In D_2O solution, since all the exchangeable protons on the bases such as protons on $-NH_2$, -NH or -OH group are deuterium exchanged, only the resonances arising from the other base protons can be observed in the pmr spectra. Because usually no distinct separate resonances arising from individual tautomeric species of the bases can be observed, it is difficult to judge from the observed resonances which tautomeric species actually exist in solution. However the broadening of some of these base proton resonances is frequently observed.

3.2 Theoretical Work .

Theoretically, there have been numerous quantum

mechanical calculations on the prediction of the energy difference among the tautomers of various bases.[48-55] Pullman and Pullman were the first to do this kind of calculation , and established a new field of so-called quantum biochemistry.[48-52] However all the theoretical calculations are semi-empirical and the results are strongly dependent on the methods and assumptions employed. For example, the earlier Hückel MO calculations of Pullman and Pullman [48] predicted that the energy difference between the amino and imino form of cytosine base is 2.08 kcal/mole in favor of the amino form , while the same kind of the calculations by Hoffman and Ladik gave the difference of 8.64 kcal/mole[53] and Löwdin obtained the result of 2.10 kcal/mole.[54] The calculations of Bodor et al.[55] employing SCF MO method showed that the energy difference between these two tautomeric forms is 2.19 kcal/mole. All of these calculations indicated that the amino form of the cytosine base is more stable.

In the case of guanine base, the deviations of the calculations due to different methods and workers are even more serious and the results are not consistent at all. Earlier Hückel calculations of Pullman et al.[48] predicted that the lactim form of the guanine base is more stable than its lactam form by the energy difference of 5.12 kcal/mole. This is obviously contrary to that which is generally accepted. A similar calculation of Hoffman

et al. also indicated that the energy difference is 12.48 kcal/mole in favor of the lactim form. However, the Hückel MO calculation of Löwdin and the recent SCF MO calculations of Bodor et al.[55] both showed that the lactam form of this base is much more stable than the lactim form, by 9.89 kcal/mole in Löwdin's work and 10.75 kcal/mole in the latter.

It should be noted that the theoretical calculations are made for the molecules in their vacuum states and hence have neglected the solvation effects. It is therefore questionable whether the results of these theoretical calculations can be extended to their situation in solution.

A summary of some of the previous important studies on the tautomerism of nucleic acid bases is given in Tables IA and IB.

4. Proton Magnetic Resonance Spectroscopy as a Tool for the Study of Tautomerism of Nucleic Acid Bases.

In spite of the above efforts, our knowledge of the tautomerism of nucleic acid bases is still fragmentary and incomplete, and in particular experimental data are still lacking on the kinetics of tautomerism. Recently we have shown that proton magnetic resonance spectroscopy is an effective method for the examination and the elucidation of the tautomerism in nucleic acid bases. It

has frequently been observed that the cytosine H_5 resonance in the pmr spectra of cytosine and related derivatives and guanine H_8 resonance in the pmr spectra of guanine nucleoside as well as nucleotides exhibited unusual broadening under certain conditions of temperature and solution pD's. We have shown that this linebroadening arises from the tautomeric exchange between the amino and imino species in the case of cytosine derivatives and between the lactam and lactim species in the case of guanine derivatives. From a detailed study of the cytosine H_5 as well as guanine H_8 linebroadening under various experimental conditions, we have been able to determine the equilibrium as well as the dynamic properties of the tautomerism in these two nucleic acid bases. From the results of our study, it turns out that the tautomerism of these bases can be quite slow under certain conditions of temperature and solution pH and the percentages of the minor tautomers estimated from this study are significantly higher than what are generally accepted. The elucidation of the controversial results of this study are presented in the subsequent chapters of this part of the thesis.

TABLE IA

Workers	Method	Compound	Energy difference between amino and imino tautomers	Tautomeric equilibrium constant
Pullman et al. (B. B. A. 64, 1962)	Hückel MO	Cytosine	ΔR = -2.08 kcal/mole (amino form more stable)	--
Hoffman etal. (Ad. Chem. Phys. 7, 1964)	Hückel MO	Cytosine	ΔR = -8.64 kcal/mole	--
Löwdin (Ad. Quant. Chem. 2, 1965)	Hückel MO	Cytosine	ΔR = -2.10 kcal/mole	--
Bodor et al. (J. A. C. S. 92, 1970)	SCF MO	Cytosine	ΔH = -2.19 kcal/mole	--
Kenner et al. (J. Chem. Soc. 1955)	UV and thermodynamic measurement	Cytosine	--	$K_T = 10^{4.7}$ (H_2O)
Brown et al. (J. Chem. Soc. 1968)	IR and UV	5, 6-dihydro-Cytosine	--	K_T = 25 (in H_2O) K_T = 0.1 (in $CHCl_3$)
Heléne et al.	UV	Cytosine	--	K_T = 30
Present work	NMR	5′-CMP	$\Delta H^\circ \sim$ -1.10 kcal/mole	$K_T \approx 6.7$ (D_2O)

ΔR = difference in resonance energy

ΔH = difference in heat of atomization

$K_T = \frac{\text{population of amino form}}{\text{population of imino form}}$

TABLE IB

Workers	Method	Compound	Energy difference between lactam and lactim tautomers	Tautomeric equilibrium constant
Pullman et al. (B. B. A. 64, 1962)	Hückel MO	Guanine	ΔR = 5.12 kcal/mole (lactim form more stable)	--
Hoffman et al. (Ad. Chem. Phys. 7, 1964)	Hückel MO	Guanine	ΔR = 12.48 kcal/mole	--
Löwdin (Ad. Quant. Chem. 2, 1965)	Hückel MO	Guanine	ΔR = -9.89 kcal/mole (lactam form more stable)	--
Bodor et al. (J. A. C. S. 92, 1970)	SCF MO	Guanine	ΔH = -10.7 kcal/mole	--
Wolfenden (J. Mol. Biol. 40, 1969)	UV study	Adenosine	--	$K_T = 10^{4.0}$ (H_2O)
Wolfenden (J. Mol. Biol. 40, 1969)	UV study	Inosine	--	K_T = 1 (H_2O)
Present work	NMR	2′-GMP	$\Delta H^\circ \sim$ -1.14 kcal/mole	$K_T \approx$ 6.8 (D_2O)

II. UNEXPLAINED BROADENING OF PROTON MAGNETIC RESONANCES IN NUCLEIC ACID BASES.

1. Unexplained Physical Phenomenon.

The proton resonances of some nucleic acid bases as well as their derivatives in aqueous solution often exhibit unusual linebroadening under certain conditions of the temperature and the solution pH. This phenomenon has been observed by many nucleic acid chemists and reported in a number of theses and papers.[56-59] In some cases this linebroadening could easily be reduced or removed after the purification of the sample, or the addition of chelating agents to the sample solution. However, there are some resonances, such as the H_5 resonance of cytosine base, the H_8 resonances of guanine base as well as the H_2 and H_8 resonances of adenine base for which linebroadening is intrinsic and is independent of the purification methods employed. For example at 100 M H_z the H_5 resonance of the cytosine base in cytosine as well as its derivatives exhibits a linewidth of 5 to 6 H_z at room temperature and pD 6.0. Similarly the H_8 resonance of some guanine nucelosides and nucleotides is ~ 8 H_z under the same experimental conditions. Although frequently observed, this phenomenon has escaped interpretation. The explanation of this interesting phenomenon is the subject of this thesis.

From the theoretical point of view, the line-broadening of the resonance in a molecular system would indicate the shortening of its transverse relaxation time (T_2). Generally speaking, the observed linewidth $\Delta\nu_{obs}$ is related to its transverse relaxation time in the normal absorption spectra by $\Delta\nu_{obs} = \frac{1}{\pi T_2'} + \frac{1}{\pi T_2}$, where $\frac{1}{\pi T_2'}$ is the contribution from the inhomogeneity of the external magnetic field, $\frac{1}{\pi T_2}$ is the intrinsic relaxation time arising from a number of relaxation mechanisms. However the former contribution usually amounts to no more than 1 H_z in most high resolution nmn experiments. The additional line-broadening of the resonances observed in some nucleic acid bases as well as their derivatives must therefore arise from significant contributions from some unexplained relaxation mechanisms. These might include (1) effect due to extensive molecular aggregation, (2) effect of scalar coupling to a chemically exchanging nucleus or a quadrupolar nucleus, (3) spin relaxation by paramagnetic species, (4) Chemical exchange effects, and (5) relaxation by spin-rotational coupling or anisotropic chemical shift.[60-63] We now discuss each of these mechanisms in detail in this chapter in order to ascertain how each mechanism might contribute to the linebroadening observed for the H_5 resonance of cytosine base and the H_8 resonance of guanine base etc.

2. Possible Mechanisms of NMR Linebroadening.

In magnetic resonance studies, there are two important relaxation times, namely the spin-lattice relaxation time and the spin-spin relaxation time. The spin-lattice relaxation time (or longitudinal relaxation time), T_1, is a measure of the life time of the nuclear spin energy levels. The spin-spin relaxation time (or transverse relaxation time), T_2, arises from the modulation of the energy levels of a nuclear spin system and is related to the resonance width by $\Delta\nu_{obs} = \Delta\nu_o + \frac{1}{\pi T_2}$ where $\Delta\nu_{obs}$ is the observed linewidth and $\Delta\nu_o$ is the contribution from the field inhomogeneity. In the following, we discuss how various relaxation mechanisms contribute to T_2.

2.1. Extensive Molecular Agrregations.

One of the most important relaxation mechanisms for nuclear spins is the magnetic dipole-dipole interaction between magnetic nuclei. The Hamiltonian for the dipole-dipole interaction H_D between two nuclei with spin angular momentum $\mathbf{I}_1$ and $\mathbf{I}_2$ respectively is given by

$$H_D = g_N^2 \beta_N^2 \, \mathbf{I}_1 \cdot \mathbf{D} \cdot \mathbf{I}_2$$

where g_N and β_N are nuclear g value and nuclear magneton respectively. $\mathbf{D}$ is the dipolar coupling tensor. Because of the tumbling of the molecules in solution the dipolar coupling tensor $\mathbf{D}$ is time dependent. Therefore H_D is also

time dependent. Based on the theory of fluctuation and the time dependent perturbation theory, the relaxation times T_1 and T_2 for a nuclear spin $\frac{1}{2}$ system due to dipole-dipole interactions between two nuclear spins can be calculated and the results are:

$$\frac{1}{T_1} = \frac{3}{20}\left(\frac{g_N^4 \beta_N^4}{\hbar^2 r^6}\right)\left(\frac{2\tau_c}{1+\omega_o^2\tau_c^2} + \frac{8\tau_c}{1+4\omega_o^2\tau_c^2}\right) \quad (1)$$

$$\frac{1}{T_2} = \frac{3}{20}\left(\frac{g_N^4 \beta_N^4}{\hbar^2 r^6}\right)\left(3\tau_c + \frac{5\tau_c}{1+\omega_o^2\tau_c^2} + \frac{2\tau_c}{1+4\omega_o^2\tau_c^2}\right) \quad (2)$$

where r is the distance between two nuclear dipoles, ω_o is the resonance frequency, τ_c is the correlation time- the characteristic time scale of the molecular motion. If we assume that the correlation time can be controlled by the Brownian motion, then $\tau_c = \frac{4\pi\eta a^3}{3\,kT}$ where a is the mean molecular radius, η is the solution viscosity and T is the absolute temperature. From Eq (1) and Eq (2), the relaxation times of a molecular system arising from the dipole-dipole interaction among the nuclear spins are closely related to the correlation time τ_c. For the molecules like nucleic acid bases as well as their derivatives in their monomeric state, τ_c is usually very short ($\leq 10^{-11}$ sec), such that $\tau_c\omega_o << 1$. In this limit T_1 can be shown to be equal to T_2. In this situation, both relaxation times are usually long and should not contribute any significant linebroadening to the observed linewidth.

For example, the relaxation times of the H_5 resonance of the monomeric cytosine base due to spin dipole-dipole interactions with the H_6 proton on the same base are estimated from Eq. (1) and Eq. (2) to be 5.0 sec . So the dipole-dipole relaxation contribution to the observed H_5 linewidth of this base is only 0.06 to o.1 Hz --an order of magnitude smaller than that from field inhomogeneity.

However, as the molecules extensively aggregate due to the stacking and/or pairing, the tumbling rate of the molecular aggregates becomes slow, because of the increase in molecular size as well as the solution viscosity. This increase of the correlation time shortens the transverse relaxation time and can give rise to significant line-broadening of the resonances if the molecular aggregation is extensive.

2.2. Effects of Scalar Coupling.

A nucleus I_1 having a scalar coupling $\hbar J \vec{I}_1 \cdot \vec{I}_2$ to another nucleus I_2 often exhibits a broad resonance spectrum, when the second nucleus has a short spin-lattice relaxation time due to some other mechanisms or when the second nucleus can undergo chemical exchange. We now discuss the effect of these time dependent processes on nucleus I_1. We shall assume in these analysis that the scalar couping is J (rad sec^{-1}), the spin-lattice relaxation time of the second nucleus is $(T_1)I_2$, and the chemical exchange time constant is τe . If $(T_1)I_2$ and τe are

long compared to 1/J the width of the energy levels of the first nucleus would be controlled by $1/T_2$ of the second nucleus, i.e., $\left(\frac{1}{T_2}\right)_{I_1} = \left(\frac{1}{T_2}\right)_{I_2}$. In this limit we observe a broad multiplet structure for the resonance of the first nucleus as a result of the scalar coupling with the second nucleus. We next consider the case where the multiplet structure has been collapsed.

(a) Scalar coupling of the first kind.

If $\frac{1}{\tau_e}$ is much longer than J, the multiplet structure disappears, each spin can give rise to a single broad resonance line and the scalar coupling $\hbar J \vec{I}_1 \cdot \vec{I}_2$ can become a mechanism of relaxation. If the chemical exchange dominates such that τ_e is shorter than T_1 of either spin, the scalar coupling constant becomes a function of time J (t). In this sutuation, the relaxation times of the first nucleus $(T_1)_{I_1}$ and $(T_2)_{I_1}$ produced by this time-dependent mechanism can be expressed as follows:

$$\left(\frac{1}{T_1}\right)_{I_1} = \frac{2J^2}{3} \frac{\tau_e}{1+(\omega_{I_1}-\omega_{I_2})^2\tau_e^2} I_2(I_2+1)$$

$$\left(\frac{1}{T_2}\right)_{I_2} = \frac{J^2}{3} I_2(I_2+1)\left\{ \tau_e + \frac{\tau_e}{1+(\omega_{I_1}-\omega_{I_2})^2\tau_e^2} \right\}$$

where ω_{I_1} and ω_{I_2} are resonance frequency of the first and the second nuclei respectively.

For example, in the case of cytosine base, the line broadening of the H_5 resonance may be produced by the

scalar coupling to the amino proton or deuterons which are in chamical exchange with water or D_2O.

(b) Scalar coupling of the second kind.

If the spin I_2 has a relaxation time due to some spin-lattice mechanism shorter than the inverse of the scalar couping constant $1/J$, then the effect of scalar coupling can be thought to be a local time-dependent magnetic field. $J\vec{I}_2(t)/\gamma_{I_1}$ produced by the spin I_2 and "seen" by the spin I_1. In this limit, the multiplet structure due to scalar couping is again collapsed and only a single broad resonance can be observed. The relaxation time of the first nucleus $(T_1)_{I_1}$ and $(T_2)_{I_1}$ due to this scalar relaxation mechanism can be expressed by the following:

$$\left(\frac{1}{T_1}\right)_{I_1} = \frac{2J^2}{3} I_2(I_2+1) \frac{(T_2)_{I_2}}{1+(\omega_{I_1}-\omega_{I_2})^2 (T_2)^2_{I_2}}$$

$$\left(\frac{1}{T_2}\right)_{I_2} = \frac{J^2}{3} I_2(I_2+1)\left\{\frac{(T_2)_{I_2}}{1+(\omega_{I_1}-\omega_{I_2})^2 (T_2)^2_{I_2}} + (T_1)_{I_2}\right\}$$

In nucleic acid bases, the second nucleus is often a nucleus other than another proton, such as ^{14}N nucleus which may have nuclear spin greater than $\frac{1}{2}$ and which has a quadrupole moment Q which can interact with the local electric field gradient q to produce spin-lattice relaxation. It can be shown that this quadrupolar

relaxation mechanism relaxes the quadrupolar nucleus by the rate

$$\left(\frac{1}{T_1}\right)_{I_2} = 3e^4q^2Q^2/80\hbar^2 \left\{ \frac{2\tau_c}{1+\omega_o^2\tau_c^2} + \frac{8\tau_c}{1+5\omega_o^2\tau_c^2} \right\}$$

where τ_c and ω_o are previously defined. In the extreme motional narrowing limit i.e., $\tau_c\omega_0 << 1$, we have $\left(\frac{1}{T_1}\right)_{I_2} = \left(\frac{1}{T_2}\right)_{I_2} = \frac{3}{8} e^4q^2Q^2\tau_c$. If $\left(\frac{1}{T_1}\right)_{I_2} > J$, but T_1 of the quadrupolar nucleus is sufficiently long and the spectrum of nucleus, can become broadened as a result of the scalar coupling to the quadrupolar nucleus. In this limit, we note that the width of the collapsed resonance decreases as the spin-relaxation of the second nucleus becomes more efficient.

2.3. Spin Relaxation by Paramagnetic Species in Solution.

Paramagnetic ions in solution have an extremely strong effect on the relaxation times of a nuclear spin system, even when they are present in low concentrations, because the magnetic fields at the nucleus produced by the unpaired electrons of the paramagnetic ions are proportional to the electron magnetic moment and are about 2000 times stronger than those produced by other nuclei (local field due to unpaired electrons ~ 10000 gauss). In addition to the relaxation effects due to dipole-dipole

interactions between the nuclear spin and unpaired electron spin of the paramagnetic ions, the contact shift interaction arising from the localization of the unpaired electron spin density in the nucleus also contributes significantly to the nuclear spin relaxation times. If we consider the relaxation effects which are caused by modulation of the contact hyperfine interaction $A\vec{I}\cdot\vec{S}$ between the nucleus I and unpaired electrons $\vec{S}$, the interaction Hamiltonian is given by $\mathcal{H}c(t) = A(t)\,\vec{S}(t)\cdot\vec{I}$. We assume that $\mathcal{H}c(t)$ fluctuates with a characteristic time scale τ_e where $\frac{1}{\tau_e} = \frac{1}{\tau_h} + \frac{1}{\tau_s}$ and, τ_h , τ_s are the exchange life time of the nucleus and the electron relaxation time respectively. Similarly the effective correlation time τ_c for the dipolar interaction is determined by the contribution from rotation, electron relaxation time and the exchange life time,i.e.,

$$\frac{1}{\tau_c} = \frac{1}{\tau_r} + \frac{1}{\tau_h} + \frac{1}{\tau_s}$$ where τ_r is the rotational correlation time.

The complete expressions of the nuclear relaxation times associated with the paramagnetic species (T_{1M} and T_{2M}) are given by the following:

$$\frac{1}{T_{1M}} = \frac{4}{30} S(S+1)\left(\frac{g^2\beta^2 g_N^2 \beta_N^2}{\hbar^2 r^6}\right)\left(3\tau_c + \frac{7\tau_c}{1+\omega_s^2\tau_c^2}\right) + \frac{2}{3}\frac{A^2}{\hbar^2}\left(\frac{\tau_e}{1+\omega_s^2\tau_e^2}\right) S(S+1) \qquad (3)$$

$$\frac{1}{T_{2M}} = \frac{4}{60} S(S+1)\left(\frac{g^2\beta^2 g_N^2 \beta_N^2}{\hbar^2 r^6}\right)\left(7\tau_c + \frac{13\tau_c}{1+\omega_s^2\tau_c^2}\right) + \frac{1}{3} S(S+1)\frac{A^2}{\hbar^2}\left(\tau_e + \frac{\tau_e}{1+\omega_s^2\tau_e^2}\right) \quad (4)$$

The first term in Eq. (3) and Eq. (4) represents the effect of dipolar hyperfine interactions, while the second term comes from the contact interaction. S is the unpaired electron spin, g, β are electron g value and Bohs magneton respectively. ω_s is the electron spin resonance frequency. A is the hyperfine coupling constant between unpaired electrons and the nucleus and is a measure of the localization of the unpaired electron spin density in the nucleus.

In case that $\omega_s\tau_e >> 1$, $\omega_s\tau_c << 1$, A is sufficiently large, the contact term can contribute appriciably to T_{2M} and gives a significant linebroadening to the nuclear resonance.

For example, in the presence of Mn^{++}, the water protons in the coordination of Mn^{++} ions suffer strong paramagnetic relaxation effects. If the exchange rate between the coordinated water and free water is sufficiently fast, so that chemical exchange between the coordinated water protons and Mn^{++} does not contribute appreciably to the observed relaxation times of water protons, the observed transverse relaxation time (T_2) can be expressed as the following

$$\frac{1}{T_2} = \frac{f}{T_{2M}} \quad \frac{1-f}{T_{2W}}$$

where T_{2M} and T_{2W} represent the relaxation times of the coordinated and free water protons respectively, f is the fraction of water protons in the coordination sphere Mn^{++}.

Paramagnetic ions are usually present as contaminants in nucleotide solutions as minor impurities, and can give rise to linebroadening of the nucleic acid resonances. For example the presence of 10^{-4}M of Mn^{++} ion in D_2O solution shortens the transverse relaxation times of both H_5 and H_6 resonances of the cytosine base as well as its derivatives to 0.04 sec at neutral pD and gives rise to about 6 to 8 H_z of linebroadening to these two resonances.

2.4. Lineroadening Due to Chemical Exchange Processes.

The lineshape and width (or transverse relaxation time) of nuclear magnetic resonances are sensitive to time-dependent processes which occur at rates of the order of the extent of the spectrum measured. These time-dependent processes may arise from the chemical exchange among distinct chemical or magnetic environments which have different resonance frequencies.

For processes slow compared with the critical rate which is determined by the chemical shift difference (Δ) between different chemical environments, spectra appear as superpositions of distinct resonances corresponding to

the individual species or environments present with line-broadening due to finite exchange life time. On the other hand, for processes rapid compared with the difference in resonance frequencies, one sees the resonance corresponding to the average of the two environments with linebroadening due to the incomplete averaging of the two chemical environments.

The effect of chemical exchange on the transverse relaxation time of the observed resonance in a system can simply be described by the suitable modifications of the Bloch phenomenological equations in a quantitative manner.

For a simple case where the chemical exchange occurs between two distinct environments A and B, the modified Bloch equations for a nuclear spin system are: (61)

$$\frac{dG_A}{dt} + \alpha_A G_A = -i\gamma H_1 M_{oA} + \tau_B^{-1} G_B - \tau_A^{-1} G_A \qquad (5)$$

$$\frac{dG_B}{dt} + \alpha_B G_B = -i\gamma H_1 M_{oB} + \tau_A^{-1} G_A - \tau_B^{-1} G_B \qquad (6)$$

where G_A and G_B are the complex moments of a nuclear spin system in the environments A and B respectively,

$\alpha_A = T_{2A}^{-1} - i(\omega_A - \omega)$, $\alpha_B = T_{2B}^{-1} - i(\omega_B - \omega)$ T_{2A} and T_{2B} are the transverse relaxation times in these two environments in the absence of exchange. ω_A and ω_B are the resonance frequencies in these two environments respectively. M_{oA} and M_{oB} are their respective equilibrium magnetization in

these two environments, γ and H_1 are nuclear gyromagnetic ratio and the reference field respectively. τ_A and τ_B are the mean life times of a nucleus staying in the environment A and B respectively,

In the steady state, i.e., $\frac{dG_A}{dt} = \frac{dG_B}{dt} = 0$, we may obtain a total complex moment $G = G_A + G_B$ by solving Eq. (5) and (6), i.e.;

$$G = i\gamma H_1 M_o \frac{\tau_A+\tau_B+\tau_A\tau_B(\alpha_A P_A+\alpha_B P_B)}{(1+\alpha_A\tau_A)\cdot(1+\alpha_B\tau_B)-1} = u+iv$$

where $M_o = M_{oA} + M_{oB}$, P_A and P_B are the relative populations in these two environments respectively. The intensity of absorption at frequency ω is then proportional to the imaginary part of the total complex moment i.e., v. Since the full expression is too complicated, we only examine its behavior and obtain expressions in the two limiting cases, where the life times τ_A and τ_B are long and short.

Fast Exchange. In this limit, τ_A and τ_B are small compared to the inverse of the chemical shift difference between the two environments $\tau_A, \tau_B << [2\pi(\omega_A-\omega_B)]^{-1}$.

The observed resonance will be centered on a mean frequency $\omega_{mean} = P_A\omega_A + P_B\omega_B$. By a proper approximation, a simplified Lorentzian expression for the absorption moment v can be obtained and the apparent transverse relaxation time T_2' can be extracted, i.e.;

$$\frac{1}{T_2'} = \frac{P_A}{T_{2A}} + \frac{P_B}{T_{2B}} + P_A^2 P_B^2 (\omega_A - \omega_B)^2 (\tau_A + \tau_B). \qquad (7)$$

Here we have assumed that the intrinsic relaxation times T_{2A} and T_{2B} are long compared to the inverse of the chemical shift difference between two environments, i.e.; T_{2A} or $T_{2B} \gg [\pi(\omega_A - \omega_B)]^{-1}$.

Slow Exchange. If the life time τ_A and τ_B are sufficiently long compared with the inverse of the chemical shift difference between two environments i.e.; τ_A or $\tau_B \gg [2\pi(\omega_A - \omega_B)]^{-1}$ the spectrum will consist of two distinct signals in the vicinity of the frequencies, ω_A and ω_B. The transverse relaxation time of each separate resonance is given by:

For resonance A $$\frac{1}{T_{2A}'} = \frac{1}{T_{2A}} + \frac{1}{\tau_A} \qquad (8a)$$

For resonance B $$\frac{1}{T_{2B}'} = \frac{1}{T_{2B}} + \frac{1}{\tau_B} \qquad (8b)$$

So, the additional linebroadening of each resonance is determined by the inverse of the exchange life time.

For a chemically exchanging nucleus which is spin coupled to a neighboring nucleus, the Bloch equations can not be used to describe its lineshape. In this situation, density matrix methods must be employed to calculate its lineshape. An example for the three-site exchange in the presence of spin-spin coupling is given in Appendix 1.

Chemical Exchange Linebroadening due to the Contact Shift Interaction of Paramagnetic Ions.[63]

We shall present a specific example to explain how paramagnetic ion contaminants sometimes give rise to chemical exchange broadening to the observed nuclear resonance. The association between the paramagnetic ion M and the molecule B to form a complex BM is typically represented by the following equation:

$$B + M \rightleftharpoons BM$$

If the resonance of the molecule suffers strong hyperfine contact interaction upon paramagnetic ion complexation, then there is large chemical shift difference $\Delta\omega_M$ between them in the free and the complexed molecules. $\Delta\omega_M$ is related to the hyperfine coupling constant A by the Bloembergen relation.[72]

$$\frac{\Delta\omega_M}{\omega_o} = 4\ I\ (I+1)\ S\ (S+1)\ A\gamma_e/9kT\gamma_N$$

where ω_o is the nuclear resonance frequency. I and S are nuclear and electron spin angular momentum respectively and γ_N and γ_e are gyromagnetic ratios of nucleus and electron respectively. If the rates of association between the molecule B and the paramagnetic ions are within the proper time scale, and the paramagnetic relaxation term of $(\frac{1}{T_{2M}})$ is not large enough to affect the observed linewidth, the chemical exchange due to paramagnetic ions does contribute appreciably to the transverse relaxation time of the

observed nuclear resonance. In this situation, we may solve a similar set of modified Bloch equations to obtain the expression for the transverse relaxation time T_{2p} of the nuclear resonance under consideration. The result is[73]

$$T_{2p}^{-1} = \tau_B^{-1} \left[\frac{1/T_{2M}^2 + 1/(T_{2M}\tau_{BM}) + \Delta\omega_M^2}{\left(\frac{1}{T_{2M}} + \frac{1}{\tau_{BM}}\right)^2 + \Delta\omega_M^2} \right]$$

Here τ_B and τ_{BM} are the mean life time of the molecule in the free and the complexed state respectively. In this derivation, we have assumed that the transverse relaxation time in the absence of the paramegnetic ions is much longer than that in the complex (T_{2M}) and the relative population of the complex (P_{BM}) is much lower than that of the free state.

We now investigate the behavior of T_{2p} in the following limiting cases[73]

a). $$\Delta\omega_M^2 >> \frac{1}{T_{2M}^2}, \frac{1}{\tau_{BM}^2}; \quad \frac{1}{T_{2p}} = \frac{P_{BM}}{\tau_{BM}} = \frac{1}{\tau_B}$$

This corresponds to the limit of slow exchange, i.e.; the contact shift $\Delta\omega_M$ is large compared to the paramagnetic relaxation rate $(\frac{1}{T_{2M}})$ as well as the inverse of the mean life time of the complex (τ_{BM}). As we expected, we find that T_{2p} is simply equal to the mean life time of that in the free state in this limit.

b). $$\frac{1}{\tau_{BM}^2} >> \Delta\omega_M^2 >> \frac{1}{(\tau_{2M}\tau_{BM})}$$

$$\frac{1}{T_{2p}} = P_{BM}\,\tau_{BM}\,\Delta\omega_M^2$$

In this limit, the equation for the fast exchange can be used to describe the relaxation time or the linewidth of the observed resonance.

c). $$\frac{1}{T_{2M}^2} >> \Delta\omega_M^2 \,,\ \frac{1}{\tau_{BM}^2}\quad \frac{1}{T_{2p}} = \frac{P_{BM}}{\tau_{BM}} = \frac{1}{\tau_B}$$

If the paramagnetic relaxation rate is fast compared to both the contact shift $\Delta\omega_M$ and the inverse of the mean life time of exchange for the complex, $\frac{1}{\tau_{BM}}$, $\frac{1}{T_{2p}}$ is simply controlled by the rate of chemical exchange.

d). $$\frac{1}{T_{2M}\tau_{BM}} >> \frac{1}{T_{2M}^2}, \Delta\omega_M^2 \,,\quad \frac{1}{T_{2p}} = \frac{P_{BM}}{T_{2M}}$$

Under this condition, $\frac{1}{T_{2M}}$ is simply determined by the paramagnetic relaxation process (T_{2M}).

Since all the parameters in the above T_{2p}^{-1} expressions are temperature dependent; a temperature as well as frequency dependence study of the linewidth of the observed nuclear resonances often enables us to distinguish between the various limits. In particular, in the

case (a) and (b), where the paramagnetic relaxation contribution to the relaxation time of the resonance is not important compared to that from chemical exchange, one merely has ordinary chemical exchange linebroadening phenomenon. At low temperatures, when the exchange rate is slow, the linewidth of the major resonance species (free species) is determined by the inverse of its exchange mean life time (τ_B). At high temperatures when the exchange rate becomes fast, one observes a field dependence of the resonance linewidth and the observed linewidth decreases with increasing temperature.

The contact shift $\Delta\omega_M$ varies with the paramagnetic ion species and the nature of the complex. In some cases $\Delta\omega_M$ can be as large as 100 ppm. So the presence of minor paramagnetic ion contaminants in the sample can also give rise to the chemical exchange linebroadening to the resonance of the sample molecules. For example, if $\Delta\omega_M$ is 100 ppm, and the binding constant between the paramagnetic ion and the molecules is sufficiently high, the contamination of 10^{-4} M of paramagnetic impurities in a 0.1 M of sample solution can give rise to an exchange linebroadening of $20H_z$ at 220 M H_z to the observed nuclear resonance.

2.5. Other Relaxation Mechanisms.

In addition to the above relaxation mechanisms which may broaden the proton resonances of nucleic acid bases, two other relaxation mechanisms exist which perhaps

should be commented on. These are (a) spin-rotational coupling, (b) modulation of the anisotropic chemical shielding by rotational tumbling. Spin-rotational interactions have been shown to provide an important relaxation mechanism for ^{19}F and ^{13}C. Except for very light molecules such as H_2 and CH_4, however, this relaxation mechanism is relatively unimportant for protons, since this spin-rotational coupling constants are usually one order of magnitude smaller than those of ^{19}F and ^{13}C. The chemical shift anisotropy mechanism usually does not contribute to the relaxation of a nuclear spin unless the shift of the anistropy is of the order of 100 ppm, and only then when the correlation time associates with the rotational coupling is very long. Since chemical shift anisotropy for proton is typically less than 10 ppm this mechanism is unimportant for protons. It is therefore possible to rule out both the spin-rotational coupling and anisotropic chemical shielding as sources of the broadening of proton resonances in the nucleic acid bases.

III. EXPERIMENTAL

1. Materials

A number of cytosine nucleosides and nucleotides were selected for this work. These compounds were obtained from various sources: cytosine and cytidine were purchased from Calbiochem, Los Angeles, Calif.; cytidine-2'-monophosphate (2'-CMP) was obtained from P-L Biochemicals, Inc., Milwaukee, Wisc.; and 3'-CMP was obtained from Sigma Chemical Co., St. Louis, Mo.; 3-methylcytidine (methosulfate salt) was obtained from Mann Research Laboratory, New York, N.Y.; 4-N,N-dimethylcytosine was synthesized from 4-ethoxyl-2-hydroxyl-pyrimidinone provided by Cyclo Chemical Corp., Los Angeles, Calif., according to the procedure of Wempen et al.;[64] cytidine-5'-triphosphate (5'-CTP) was also purchased from this company. 5'-CMP, which was used as a model compound in our kinetic studies, was purchased from the following three sources: (1) Sigma Chemical Co., (2) Calbiochem, and (3) Cyclo Chemical Corp.

The following guanine nucleosides and nucleotides were obtained either in the free acid or sodium salt form: guanosine-2'-monophosphate (2'-GMP), 3'-GMP and 5'-GMP were purchased either from Sigma Chemical Co., St. Louis, Mo., or from Calbiochem, Los Angeles, Calif., as grade A reagents. Guanosine and 2-N,N-dimethylamino-methylene-guanosine were obtained from Calbiochem. 1-methyl-guanosine and 6-methoxy-

purine riboside were purchased from Cyclo Chemical Corp., Los Angeles, Calif.

The cytosine H_5 resonance of various derivatives was found to vary with the source of the sample as well as the method of purification indicating the presence of contaminants. The variation of the H_5 linewidth with sample source in the case of 5'-CMP is summarized in Table Ic. However, after the original sample was recrystallized in water-ethanol mixture[16] and then passed through a Dowex 50W-X8 cation exchanger which was washed thoroughly with EDTA before use, the linewidth was reproducible from sample to sample (Table I). The absence of paramagnetic impurities Mn^{+2}, Ni^{+2}, Co^{+2}, Cu^{+2} at significant levels after purification was confirmed by atomic absorption analysis of the sample by Schwarzkope Microanalytical Laboratory, Woodside, N.Y.

The experimental results reported in this work have been obtained on samples purified in the above manner.

2. Methods

Sample solutions were prepared at concentrations of 0.01 to 0.05 M in 99.7% D_2O supplied by Columbia Organic Chemicals, Columbia, S.C. Some of the samples were lyophilized from D_2O prior to sample preparation, in order to reduce the magnitude of the HDO peak in the spectra. The pD of each solution was measured with a Leeds and Northrup 7401 pH meter, equipped with miniature electrodes, and

TABLE IC.

Variation of the cytosine H_5 resonance of 5′-CMP with sample source and sample treatment

No.	Sample Source	History	Linewidth of H_5 resonance[a] Without Purification	Cation Exchange(I)[b]	Recrystallization	Cation Exchange(II)
1	Calbiochem	Purchased as free acid in Nov. 1963. Lot No. 540671.	11 ± 1 Hz			
2	Sigma	Purchased as disodium salt in Dec. 1968. Lot No. 124B-0260.	11 ± 1 Hz			
3	Calbiochem	Purchased as free acid in Jan. 1969. Lot No. 540671.	11 ± 1 Hz			
4	Calbiochem	Purchased as free acid in Feb. 1970. Lot No. 540671.	2.5 Hz	8 ± 1 Hz	11 ± 1 Hz	11 ± 1 Hz
5	Sigma	Purchased as free acid in Feb. 1970. Lot No. 125B-7310.	2.5 Hz	9 ± 1 Hz	11 ± 1 Hz	11 ± 1 Hz
6	Cyclo	Purchased as disodium salt in March 1970. Lot No. H3793J.	10 ± 1 Hz			11 ± 1 Hz

[a] Measured at 220 MHz at 16°C and pD 5.8 ± 0.1.

[b] Cation Exchange(I): the resin was not washed with EDTA prior to use.
Cation Exchange(II): the resin used was previously washed with EDTA.

was calculated from the observed pH meter reading plus 0.4 (the standard correction).[65] pD adjustments were made with either 1 M DCl or 1M NaOD.

The pmr spectra of the various nucleoside and nucleotide solutions were recorded on Varian HA-100 and HR-220 nmr spectrometers. A C-1024 time-average computer was used to enhance the signal-to-noise ratio. Chemical shifts were measured relative to an external TMS capillary, which also provided the field/frequence lock signal for HA-100 operation. The error in the measurement of the chemical shifts and linewidths is ± 0.5 Hz. The ambient probe temperature was 30 ±1°C in the case of the HA-100 spectrometer and 17 ±1°C for the HR-220. In the variable-temperature studies, the sample temperature was controlled to ± 1°C by means of a variable temperature controller, and was measured using the methanol and ethylene glycol standards together with the calibration curves supplied by Varian.

IV. RESULTS

1. Cytosine.

The pmr spectra of cytosine and many of its nucleoside and nucleotide derivatives have previously been reported and discussed.(56,57,66) In our present work, we shall primarily be concerned with the spectral behavior of the H_5 and H_6 resonances of the cytosine base. These resonances appear as doublets due to the spin-spin coupling between these protons and are readily identified by the spin-spin splitting of 8 Hz. The H_5 doublet is found 6 ppm downfield from TMS, and the H_6 doublet 8 ppm downfield from TMS. The H_6 resonances are always well-defined and relatively sharp; by contrast, the H_5 resonances are frequently broadened, often so much as to escape observation. This H_5 line broadening is both pD and temperature dependent. The elucidation of this previously unexplained linebroadening is the principal objective of this work.

1.1. pD studies.

As illustrated in Figures 4 and 5 , the spectral positions of the H_5 and H_6 resonances of cytosine and its nucleoside and nucleotide derivatives are pD dependent. The chemical shifts of these protons, for example, change abruptly in the pD region 4 to 5, where ring protonation is known to occur. Above pD 5, however, the chemical shifts are relatively independent of pD, except in the case

Figure 4. pD dependence of the chemical shifts of the H_5 resonances of cytosine and related compounds at 30°C.

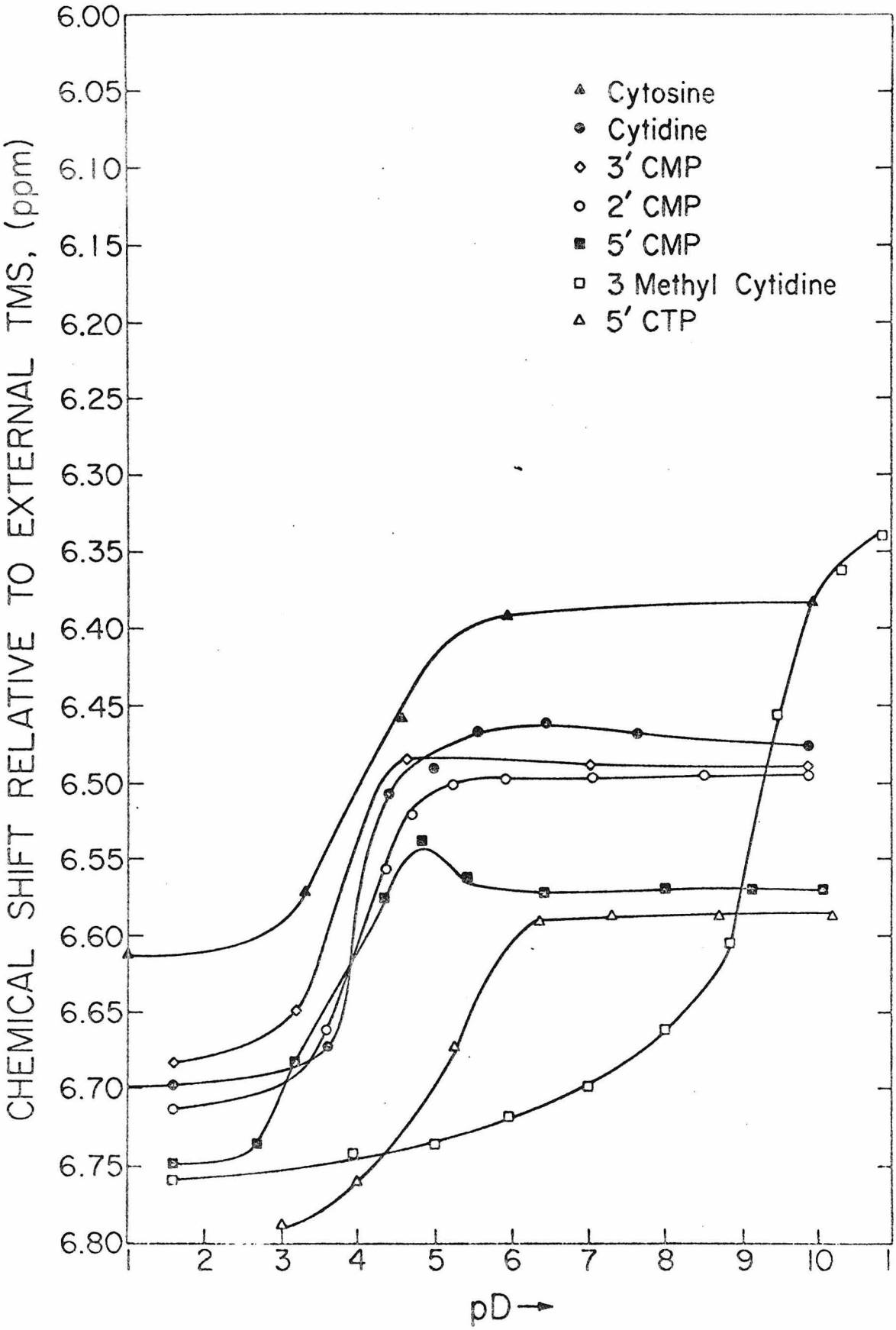
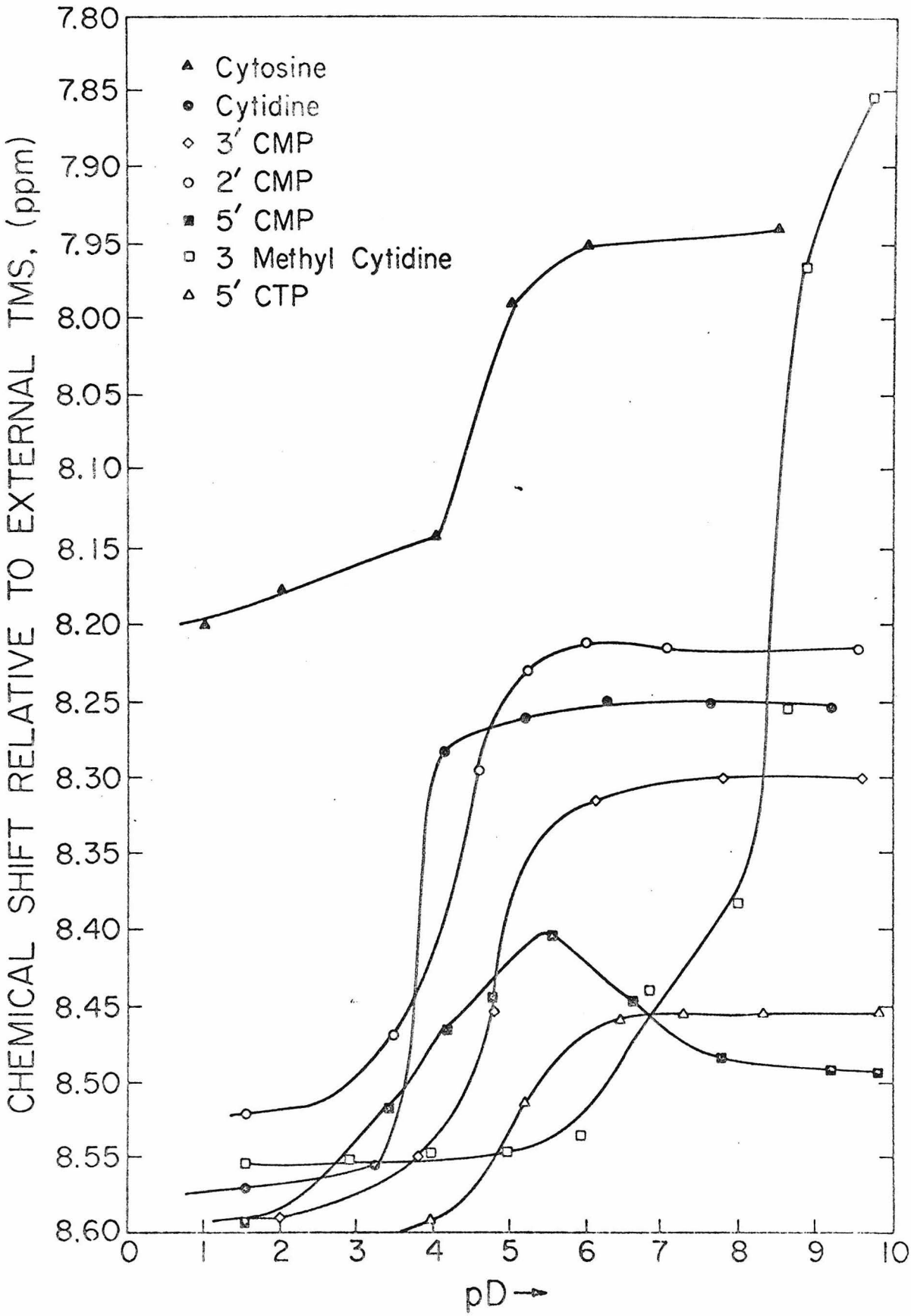

Cytosine
Cytidine
3′ CMP
2′ CMP
5′ CMP
3 Methyl Cytidine
5′ CTP
CHEMICAL SHIFT RELATIVE TO EXTERNAL TMS, (ppm)
6.00
6.05
6.10
6.15
6.20
6.25
6.30
6.35
6.40
6.45
6.50
6.55
6.60
6.65
6.70
6.75
6.80
1
2
3
4
5
6
7
8
9
10
11
pD→

Figure 5. pD dependence of the chemical shifts of the H_6 resonances of cytosine and related compounds at 30°C.


Cytosine
Cytidine
3′ CMP
2′ CMP
5′ CMP
3 Methyl Cytidine
5′ CTP
CHEMICAL SHIFT RELATIVE TO EXTERNAL TMS, (ppm)
7.80
7.85
7.90
7.95
8.00
8.05
8.10
8.15
8.20
8.25
8.30
8.35
8.40
8.45
8.50
8.55
8.60
0
1
2
3
4
5
6
7
8
9
10
pD →

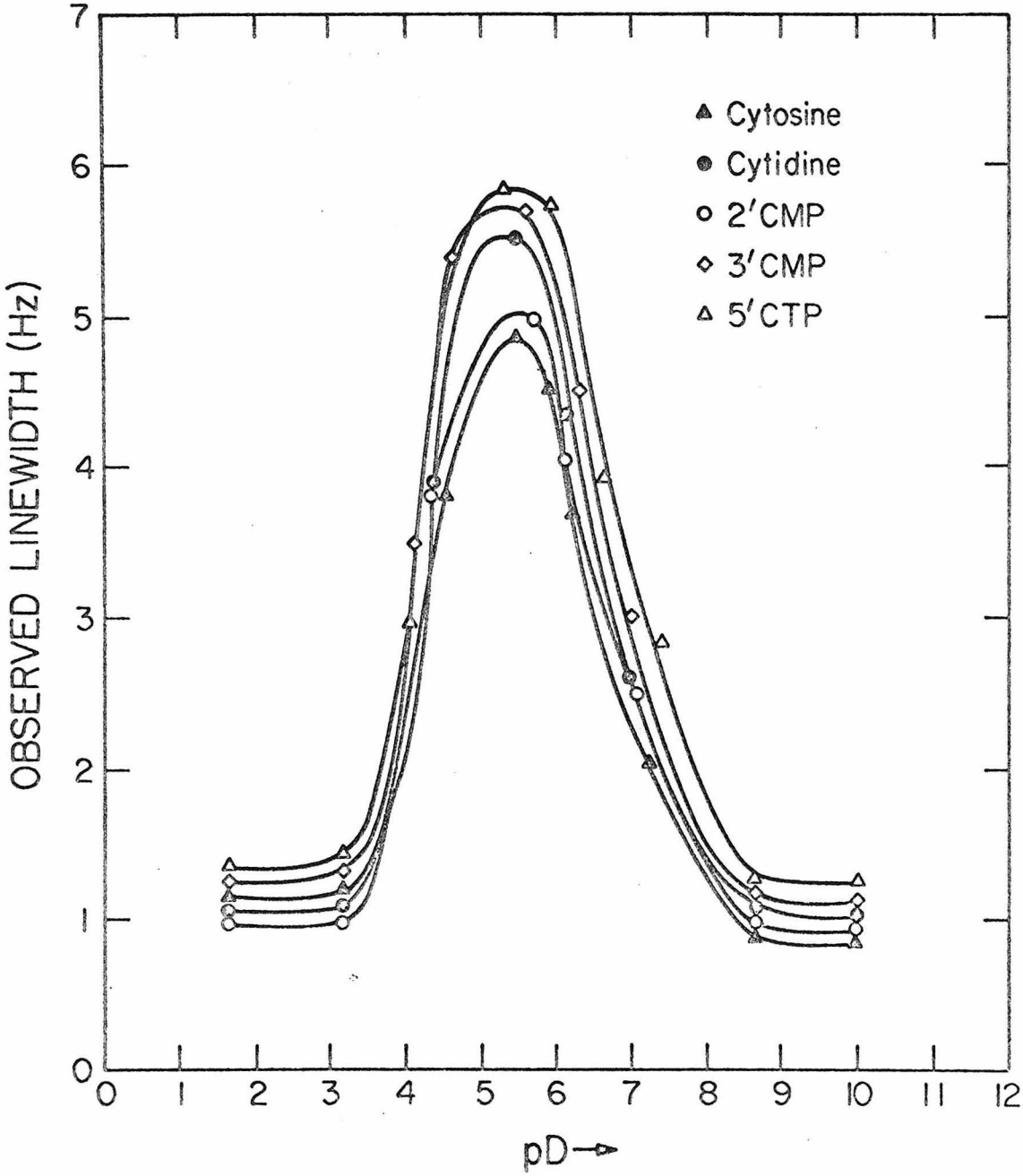
5'-CMP, where effects due to ionization of the phosphate group at pD 6 are apparent. The phosphate group ionization is important, since in pyrimidine nucleotides, the conformation of the base about the glycosidic bond is expected to be predominantly anti,[56,66,67] and in this base conformation, the phosphate group is in juxtaposition with that part of the cytosine base bearing the H_5 and H_6 protons. As expected, the deshielding effect of the negatively charged phosphate group is more pronounced for the H_6 than for the H_5 proton because of the closer proximity of the phosphate group to the H_6 proton.

In contrast to the essentially constant chemical shifts, the linewidth of H_5, resonance is strongly pD dependent in the pD range 5 to 7, exhibiting a maximum near pD 6. The width of the H_6 resonance is also pD dependent, but to a much lesser extent. The H_1, ribose resonance remains sharp throughout the pD range investigated.

As shown in Figs. 6 and 19 , the extent of broadening of the H_5 resonance, as well as the pD range over which the broadening is important, depends slightly on the compound examined. We shall, therefore, discuss each case separately.

(a) Cytosine. In this case, the linewidth of the H_5 resonance increases abruptly in the pD range 3 to 5, reaching a maximum of 4.8 Hz at pD 5.5, and then decreasing to a width of 1.5 Hz at pD's above 8.0. Over the same pD

Figure 6. pD dependence of the linewidths of the H_5 resonances of the cytosine and related compounds at 30°C and 100 MHz.


Cytosine
Cytidine
2′CMP
3′CMP
5′CTP
OBSERVED LINEWIDTH (Hz)
pD→
0
1
2
3
4
5
6
7
8
9
10
11
12

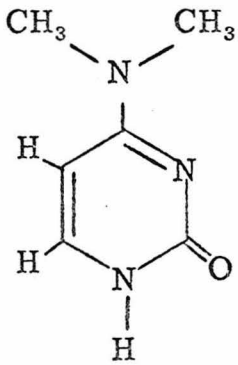
range, the H_6 resonance behaves qualitatively the same but reaches a maximum of only 1.4 Hz at pD 5.0.

(b) Cytidine. In spite of the addition of the ribose moiety, the pD dependence of the linewidth of both H_5 and H_6 resonances is similar to that of cytosine as can be seen from Figure 6 . The maximum H_5 linewidth is about 5.5 Hz at pD 5.8. The H_6 broadening is again significantly less and the H_1' resonance does not appear to broaden.

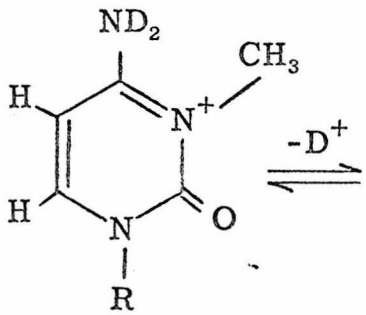
(c) 5'-CMP, 3'-CMP, 2'-CMP. The addition of the phosphate group to the nucleoside has little effect on the qualitative behavior of the H_5 linewidth. In the case of 5'-CMP, the maximum H_5 linewidth observed is 6.0 Hz at pD 6.0. In 3'-CMP the maximum H_5 linewidth is about 5.8 Hz at pD 6, while in 2'-CMP the maximum width is 5 Hz. In these nucleotides, the H_6 resonance is again broadened but to a much lesser extent than the corresponding H_5 resonance.

(d) 2'-deoxycytidine, 5'-dCMP, CpC. Similar line broadening of the H_5 resonance of the cytosine base was also observed in these compounds. The absence of the 2'-OH group on the sugar moiety in the 2'-deoxynucleoside and 2'-deoxynucleotide as well as the base-stacking interaction between the two cytosine bases in the dinucleotide, CpC, apparently has no significant effect on the H_5 line broadening.

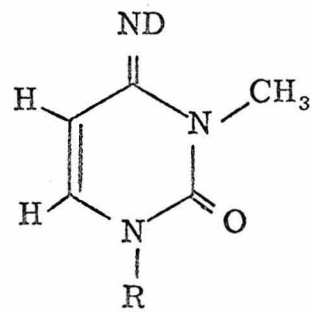
(e) 4-N,N-dimethylcytosine, 3-methylcytidine methosulfate. For these two cytosine derivatives, both the H_5 and H_6 resonances were observed



4-N, N-dimethyl-cytosine



3-methyl-cytidine



$7.5 < pD < 9.5$

3-methyl, 4-imino cytidine

to be sharp throughout the pD range of interest. Here only one tautomeric structure can exist, the amino form in the case of 4-N,N-dimethylcytosine, and the imino form in the case of 3-methylcytidine. Although the linewidth was studied from pD 2 to 12 for comparative purposes, only the range in which the neutral tautomeric form exists is of interest here. In the case of 4-N,N-dimethylcytosine, this range is essentially that of normal cytosine base (pD 4 to 10). However, in the case of 3-methylcytidine, the neutral species exists only in a narrow pD range beginning around pD 8.0, at which the amino proton is lost to form an imino tautomer, and ending near pD 10.0, where the second amino proton is lost. The existence of two pK's involved in this latter titration process is apparent from the width of the pD titration curve. The pK values of 8.2 and 10.5 deduced from the pH titration curve were confirmed by analysis of uv spectral data.

1.2. Effects of Temperature and Magnetic Field.

We have investigated the temperature and frequency dependence of the H_5 linewidth in several of the cytosine derivatives. These studies were made in aqueous D_2O solutions as well as solutions containing high concentrations of salt (NaCl, LiCl, CsCl) at pD 5.8 ± 0.2 over the temperature range -25°C to 90°C at both 100 MHz and 220 MHz. The use of electrolytes permitted the temperature range to be extended below ~0°C. The observations were

affected somewhat by the added salt, but for a given salt concentration and at a given temperature, there did not appear to be a noticeable linewidth dependence on the nature of the electrolyte.

The results for 5'-CMP in 5 N NaCl are summarized in Figures 7 and 8 . For sake of discussion, we have divided the temperature range into three regions. In the high temperature region ($> 20^{\circ}C$), the H_5 resonance of 5'-CMP was observed to be strongly temperature and field dependent. The H_5 linewidth is much broader at 220 MHz than at 100 MHz and decreases gradually with increasing temperature. In the low temperature region ($< 0^{\circ}C$), the H_5 linewidth becomes essentially independent of field and decreases with decreasing temperature. The greatest H_5 line broadening was observed in the intermediate temperature region, where the broadening was so severe that the H_5 doublet was frequently unresolved. At 220 MHz, the maximum H_5 linewidth is 11.5 Hz at $17^{\circ}C$ and at 100 MHz, the H_5 resonance reaches a maximum width of 7.5 Hz at $5^{\circ}C$.

In conjunction with the temperature studies, we have also noted a progressive downfield shift of the H_5 resonance relative to that of H_1' as the temperature was increased. For 5'-CMP in 5 N NaCl, this differential shift, $\Delta(\delta_{H_5} - \delta_{H_{1'}})$, amounted to -2.1 ± 0.5 Hz at 100 MHz and -4.0 ± 0.5 Hz at 220 MHz over the temperature range $-20^{\circ}C$ to $50^{\circ}C$. These observations may be contrasted with

Figure 7. 100 MHz and 220 MHz pmr spectra of the H_5 resonance of 5' CMP at various temperatures.

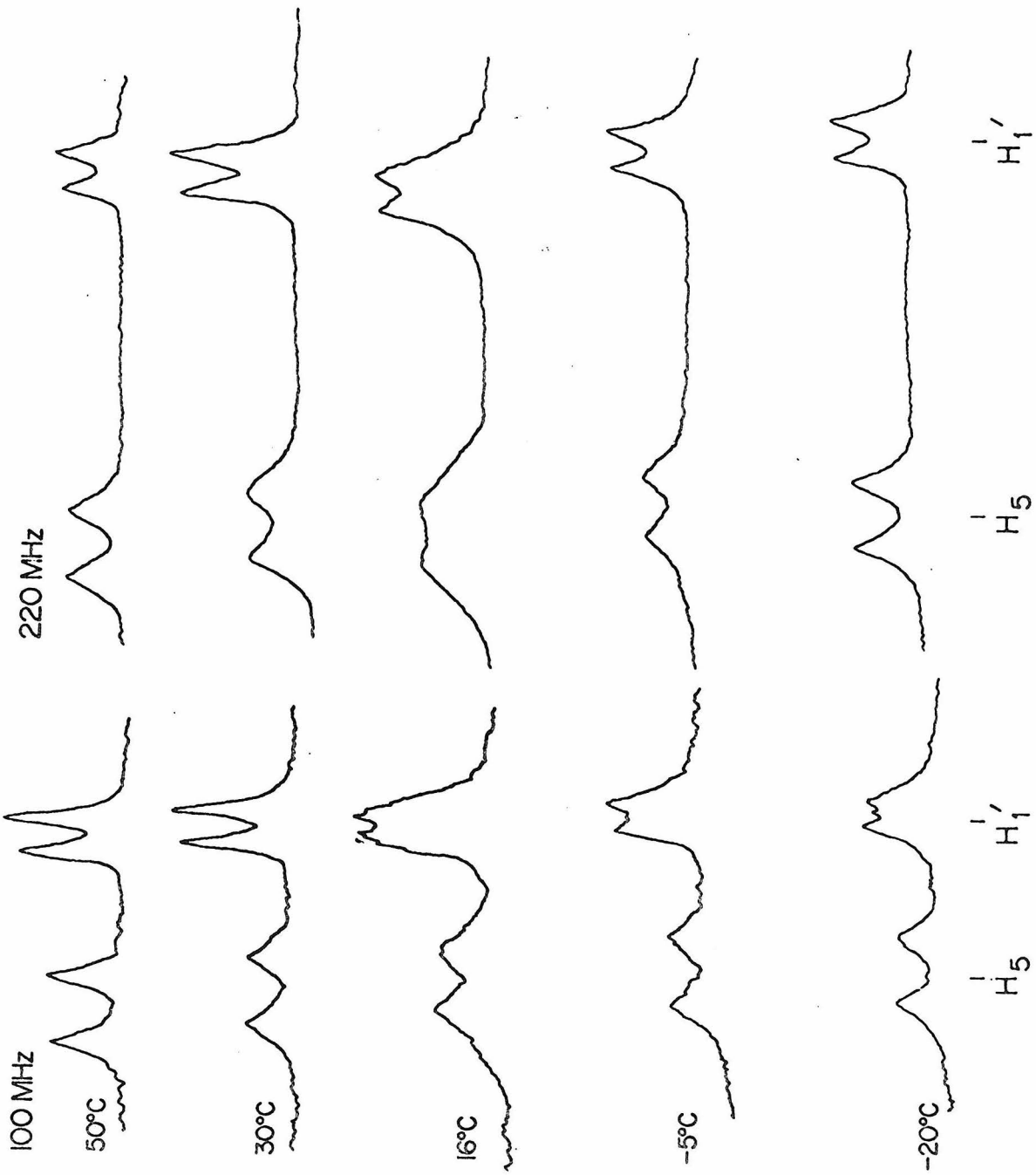
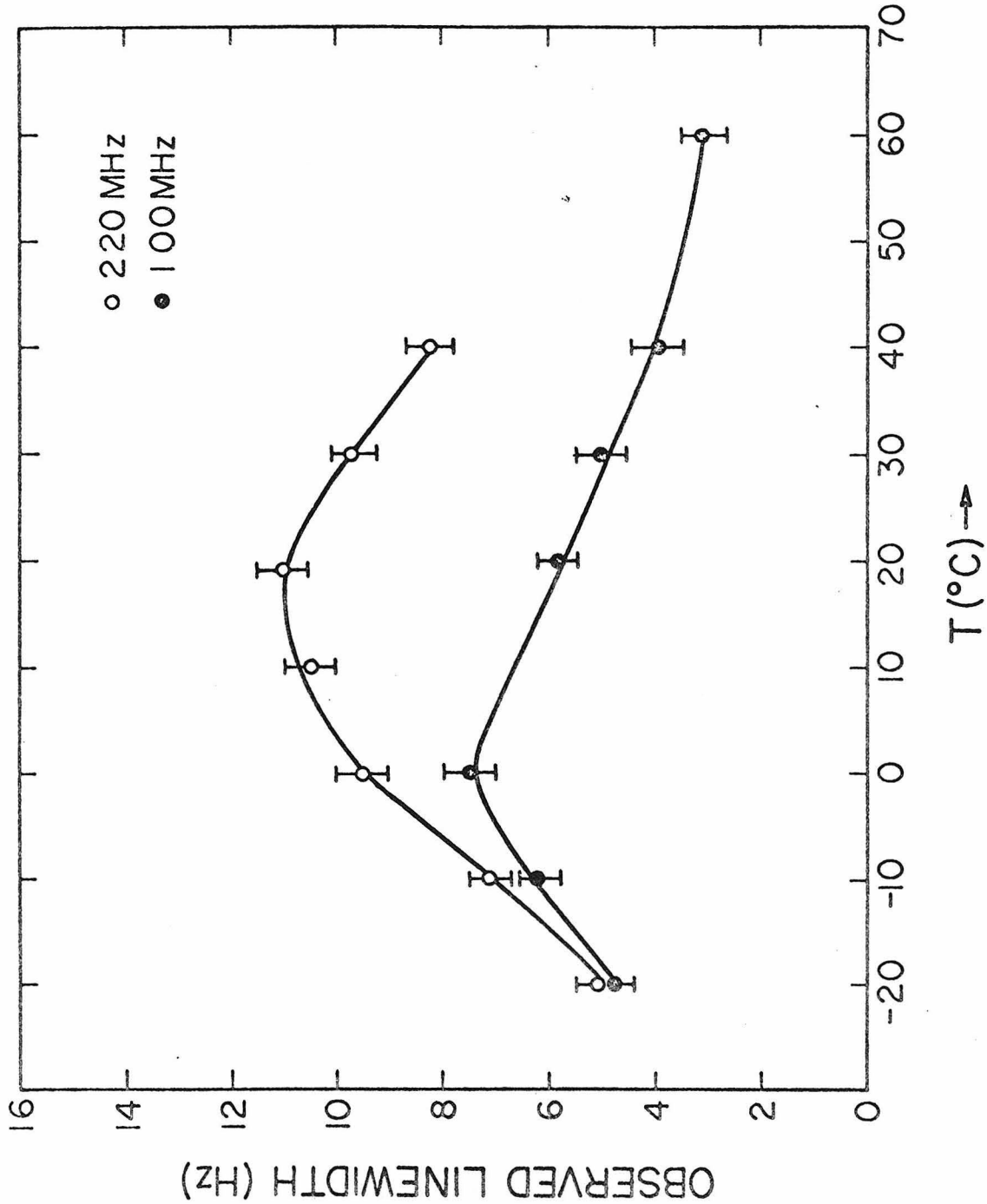

100 MHz
220 MHz
50°C
30°C
16°C
-5°C
-20°C
H5'
H1'
H5'
H1'

Figure 8. Frequency and temperature dependence of the linewidth of the H_5 resonance of a 0.02 M 5'-CMP solution containing 5 M NaCl in D_2O at pD 6.0.


o 220 MHz
• 100 MHz
OBSERVED LINEWIDTH (Hz)
T (°C) →
0
2
4
6
8
10
12
14
16
-20
-10
0
10
20
30
40
50
60
70

the almost zero differential shifts observed for the H_5 and H_1, resonances of uridine-5'-monophosphate.[68]

1.3. Intensity Measurements.

Experimentally we have observed that the intensity ratio of the H_5 and H_1' doublets changes with temperature. The results of these intensity measurements for 5'-CMP are summarized in Table II. Above 20°C, the ratio of the H_5 and H_1' intensities was found to be unity within experimental error. At low temperatures (-10°C to -25°C), this intensity ratio is reduced to 0.90 ± 0.03 at 100 MHz, and 0.93 ± 0.02 at 220 MHz. Two methods of obtaining these intensity ratios were used: spectral area measurement using a planimeter and comparison of the experimental spectra with those simulated on a digital computer.(Fig. 9)

2. Guanine.

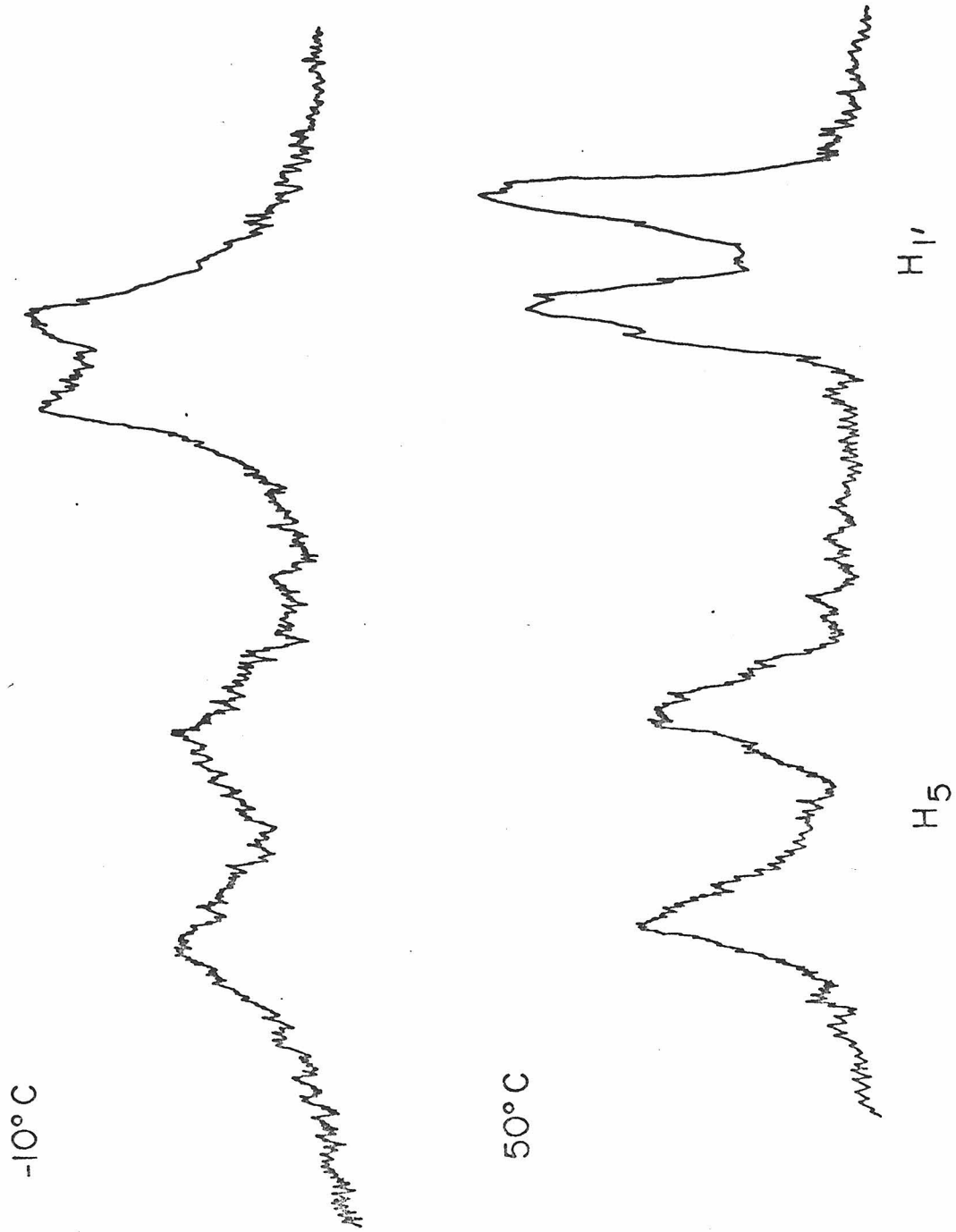
The solution properties of guanosine and its related derivatives have been studied by Gellert et al.,[69] and Miles et al.,[70] using uv, ir, optical rotation as well as x-ray crystallography, and it has been found that some of these guanine derivatives aggregate in aqueous solution to form viscous gels at intermediate pH's around room temperature for sufficiently concentrated solutions. However, this tendency towards gel formation varies greatly with the compound. 3'-GMP and 5'-GMP have been shown to form viscous gels even at concentrations of the order of

TABLE II

Intensity Measurements of H_5 and $H_{1'}$ Resonances at Different Temperatures

Field, Temperature		Intensity Arbitrary Units H_5	$H_{1'}$	Intensity Ratio	Average
HA-100	30°C	148	141	105 %	
	30°C	55	57	96.5 %	
	30°C	99	102	97.0 %	98.6 ± 3.1 %
	50°C	78	81	96.3 %	
	50°C	68	69	98.6 %	
HA-100	-10°C	126	139	90.64%	
	-10°C	102	118	86.44%	
	-20°C	96	110	92.30%	89.79 ± 3.1 %
	-15°C	62	69	89.85%	
HR-220	-20°C	48.3	50.70	95.26%	
	-20°C	94.08	100.27	93.8 %	
	-20°C	64.10	69.30	92.49%	93.06 ± 2.09 %
	-20°C	22.78	25.35	89.95%	
	-20°C	13.22	14.06	94.0 %	

Figure 9. Comparisons of the spectral area between the H_5 and H_1' resonances of 5'-CMP at -10°C and 50°C.


-10° C
50° C
H5
H1'

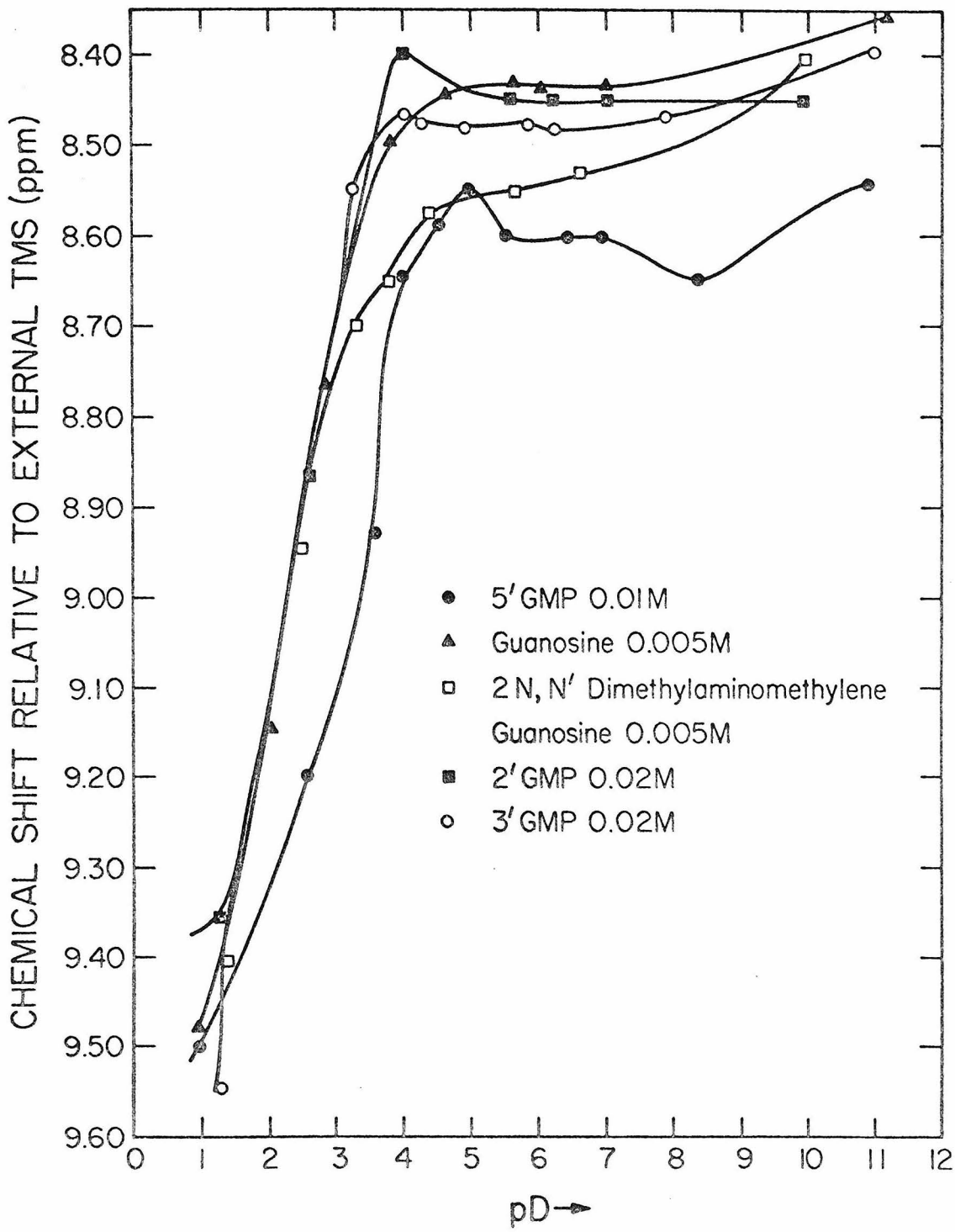
10 mM whereas 2'-GMP and 2'-deoxyguanosine exhibit little tendency towards gel formation. This strong intermolecular association (pairing and stacking) can complicate the analysis of our nmr data, but if the data were obtained under conditions where gel formation is suppressed, these complications can be circumvented.

The pmr spectra of guanosine and several of its derivatives have been previously reported and discussed.(66) The guanine H_8 resonance appears approximately 8 to 9 ppm downfield from TMS. It is a singlet but it is often broadened, even at sufficiently low nucleoside or nucleotide concentrations where gel formation should no longer prevail and for those guanine derivatives which exhibit little tendency towards gel formation. The elucidation of this linebroadening is the principal objective of this work.

2.1. pD and Concentration Studies.

The variation of the chemical shift of the H_8 proton of guanosine and its derivatives with pH in aqueous solution has been previously reported.(66) A significant downfield shift of about 1 ppm occurs between pD 2.0 to 4.0 as the result of the protonation of the base (Fig. 10). Above pD 4.0 the chemical shift of the H_8 resonance of these compounds remains essentially constant, except in the case of 5'-GMP, where effects due to the ionization of the phosphate group are apparent, indicating that the base orientation is preferentially anti in this nucleotide.

Figure 10. pD dependence of the chemical shifts of the H_8 resonance of guanosine and its derivatives.


CHEMICAL SHIFT RELATIVE TO EXTERNAL TMS (ppm)
8.40
8.50
8.60
8.70
8.80
8.90
9.00
9.10
9.20
9.30
9.40
9.50
9.60
0
1
2
3
4
5
6
7
8
9
10
11
12
pD→
5′GMP 0.01M
Guanosine 0.005M
2 N, N′ Dimethylaminomethylene Guanosine 0.005M
2′GMP 0.02M
3′GMP 0.02M

By contrast, the linewidth of the guanine H_8 resonance in many guanine derivatives is broad over the pD range 2 to 8. In some of these derivatives, it is also strongly concentration dependent. The extent of broadening as well as the concentration and pD range over which the broadening is important is strongly dependent on the compound. We therefore discuss each case separately.

(a) 2'-GMP, 2-N,N-dimethylamino-methylene-guanosine and adenylyl (3'-5') guanosine (ApG). It is known that these compounds do not undergo gel formation. Accordingly, the linewidths of the guanine H_8 resonances in these compounds were found to be concentration independent.

The pD dependence of the guanine H_8 resonance width of 2'-GMP at the concentration of 0.01 M is shown in Fig. 10 . The H_8 linewidth can be seen to be strongly pD dependent. In the low pD region, the linewidth increases with increasing pD, until it reaches a maximum of 9 Hz at pD ~ 5.0, beyond which it narrows with increasing solution pD until it reaches the limiting linewidth of ~ 1 Hz at at pD > 8.0. Insofar as we were able to ascertain, these observations are independent of the concentration over a 100-fold concentration from 0.001 to 0.1 M. Similar observations were noted for 2-N,N-dimethylamino-methylene-guanosine and the dinucleotide ApG. In the latter case, both the H_2 and H_8 resonances of the adenine residue remain sharp (2~3 Hz) throughout the whole pD range over

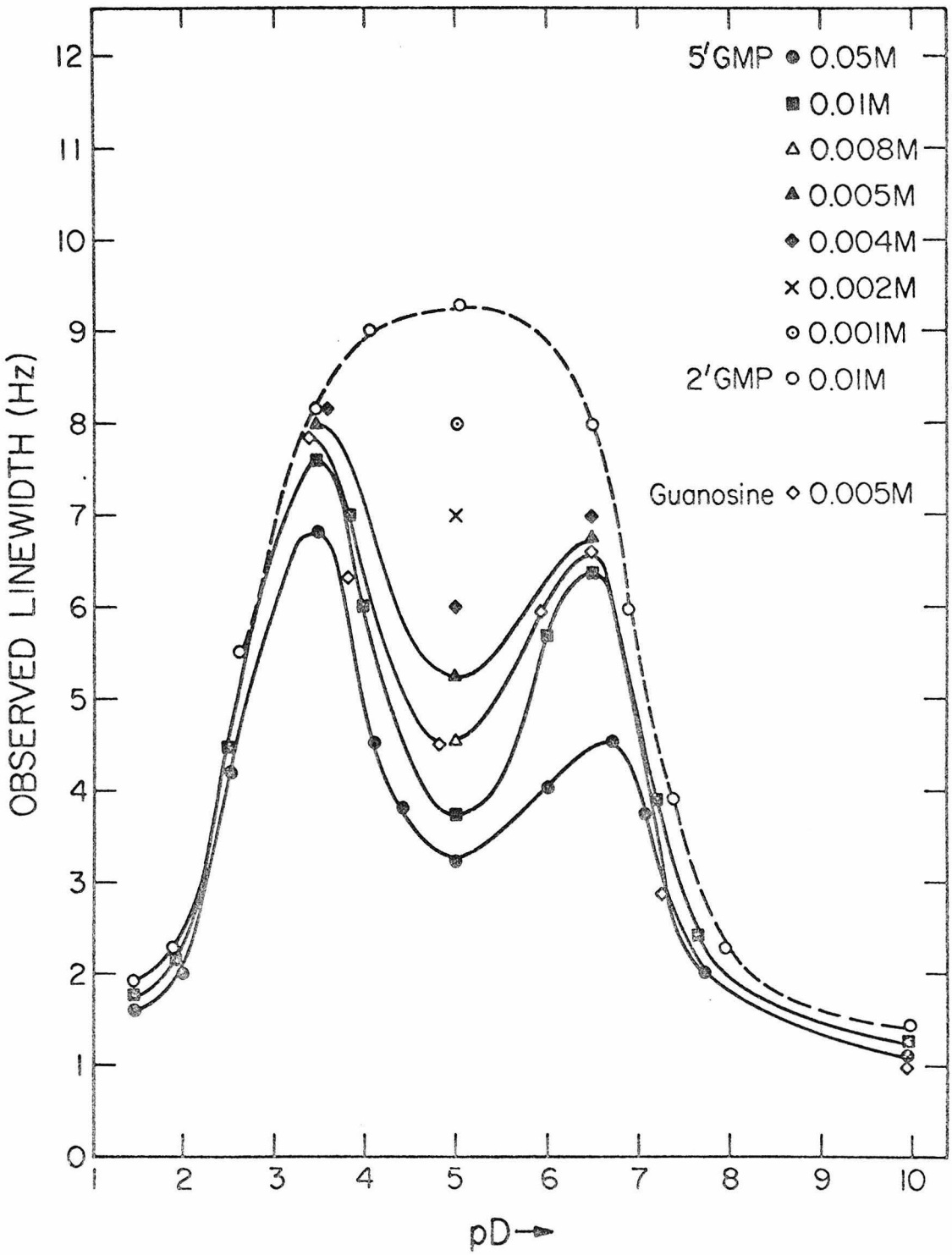
which the linebroadening of the H_8 resonance of the guanine residue was observed.

(b) Guanosine, 5'-GMP and 3'-GMP. 5'-GMP and 3'-GMP are known to form viscous gel even at low concentrations. These solutions are often viscous even at concentrations as low as 5 mM, the concentration level which was normally used in our nmr measurements.

The pD dependence of the H_8 resonance of guanosine at a concentration of 0.005 M is shown in Fig. 11 . This concentration corresponds to the solubility limit of this compound. As in the case of 2'-GMP, the spectral width of the H_8 resonance is strongly pD dependent. In fact, except in the pD range 4 to 6, the linewidth behaviour is quite similar. In the low pD region, the linewidth increases with increasing pD until it reaches a maximum width of 8 Hz at pD 3.2. Beyond this pD, it decreases abruptly to a minimum width of 3.8 Hz at pD 5.0. As the pD of the solution is further increased, the H_8 linewidth increases again to a maximum of 6.5 Hz at pD 6.5, and at still higher pD's, it narrows again with increasing solution pD until it approaches the limiting linewidth of $\sim$ 1 Hz at pD $>$ 8.0.

The linewidth behavior for the H_8 resonance of 5'-GMP at various pD's is similar to that for guanosine. For a 0.01 M solution of 5'-GMP, the two linewidth maxima were found to be 7.8 Hz and 5.8 Hz at pD 3.5 and 6.5 respectively, while the linewidth minimum at pD$\sim$5.0 is

Figure 11. pD and concentration dependence of the linewidths of the H_8 resonance of guanosine and its related derivatives at 30°C (100 MHz).


5'GMP ● 0.05M
■ 0.01M
△ 0.008M
▲ 0.005M
◆ 0.004M
× 0.002M
⊙ 0.001M
2'GMP ○ 0.01M
Guanosine ◇ 0.005M
OBSERVED LINEWIDTH (Hz)
pD→
0
1
2
3
4
5
6
7
8
9
10
11
12

3.5 Hz. However, in contrast to the conspicuous lack of dependence of the linewidth on the concentration in the case of 2'-GMP, the linewidth of the H_8 resonance here is strongly concentration dependent, particularly at low nucleotide concentrations and in the intermediate pD region 3 to 6. This behavior is also true in the case of guanosine although we did not study the concentration dependence here in detail. As shown in Fig. 11 ,at pD 5.0 the linewidth of the H_8 resonance of 5'-GMP increases gradually from 3.0 Hz at the concentration of 0.05 M to the limit of about 9.0 Hz for concentrations lower than 0.001 M. At pD 3.5, the H_8 linewidth increases from 6.0 Hz to 8.0 Hz as the concentration is decreased from 0.05 M to 0.008 M, but then remains unchanged as the solution is further diluted. At pD 6.5, the observed linewidth changes from 4.5 to 7 Hz as the concentration is decreased from 0.05 M to 0.005 M. Below this latter concentration level, the linewidth remains essentially constant at $\backsim$ 7 Hz. Despite these dramatic changes in the H_8 linewidth, there was no noticeable change in the spectral position of this resonance over the tenfold concentration range of 0.005-0.05 M.

With 3'-GMP, gel formation occurs only at concentrations in excess of 0.1 M. At the concentration levels normally used in our nmr experiment ($<$0.005 M), the linewidth behavior of the H_8 resonance of 3'-GMP is similar to that observed for 2'-GMP.

(c) 2'-deoxyguanosine, 5'-dGMP, 5'-GTP. Although these compounds do not form viscous gels in aqueous solution, the linewidth behavior of the H_8 resonance in these compounds is similar to those observed for guanosine or 5'-GMP over the same pD range. The H_8 linewidth minimum around pD 5.0 is, however, somewhat greater in these guanine derivatives than for 5'-GMP at the same concentration.

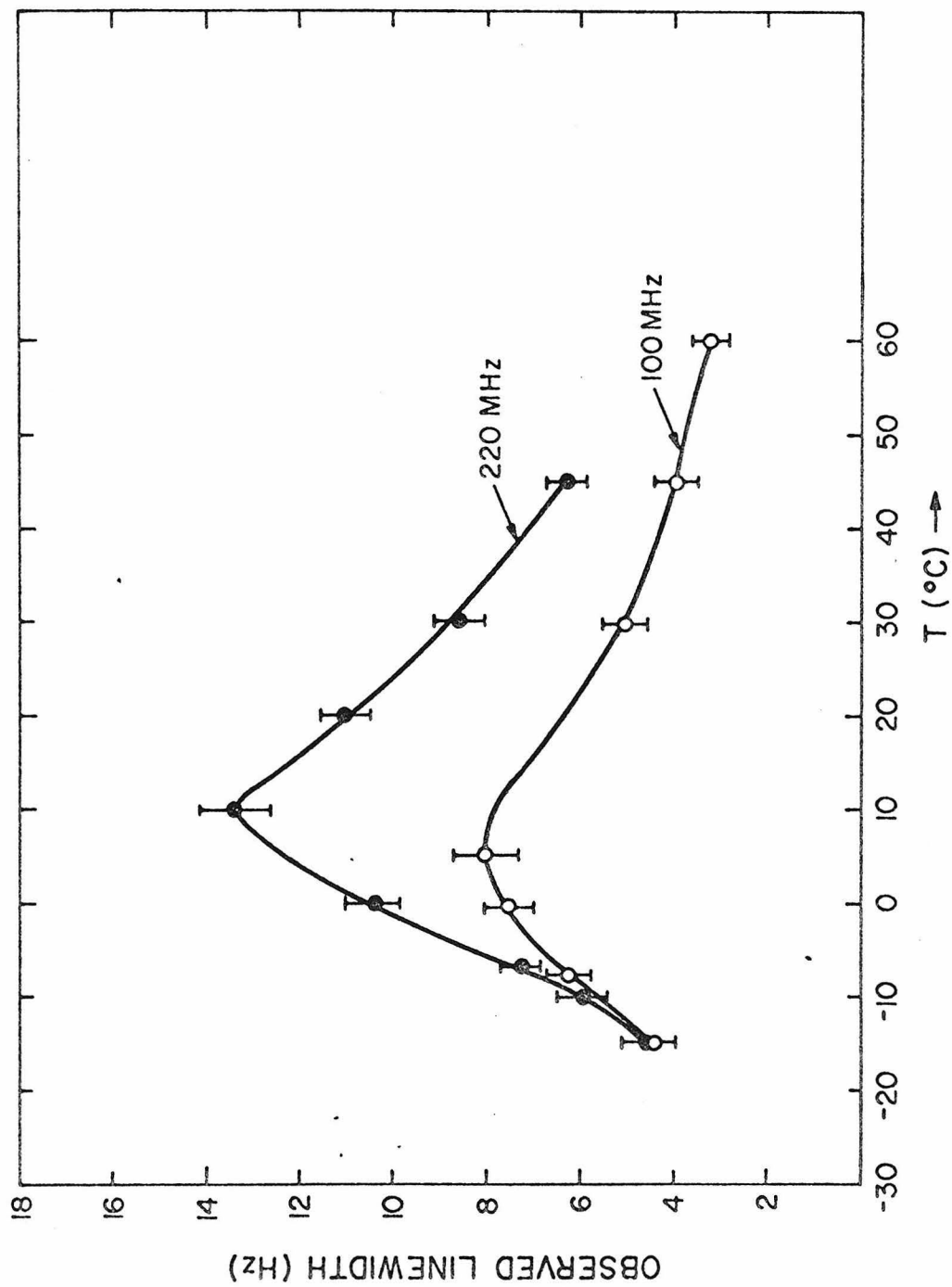
(d) 1-methyl-guanosine and 6-methoxy-purine riboside. For these two compounds, the H_8 resonances remain sharp throughout the whole pD range of interest. In these methylated derivatives, the guanine base is frozen in the lactam form in the case of 1-methyl-guanosine, and the lactim form in the case of 6-methoxy-purine riboside.

2.2. Effects of Temperature and Magnetic Field.

We have investigated the effect of temperature and magnetic field (or frequency) on the H_8 linewidth in several of the guanine derivatives.

The results for a 0.05 M 2'-GMP solution containing saturated sodium chloride at pD 7.0 $\pm$ 0.2 are summarized in Fig.12 . For convenience of discussion, we have divided the whole temperature range into three regions. In the high temperature region ($> 30^{\circ}C$), the H_8 resonance of 2'-GMP was found to be strongly temperature and field dependent. The H_8 linewidth is much broader at 220 MHz than at 100 MHz and decreases gradually with increasing temperature. In the low temperature region ($< 0^{\circ}C$), the

Figure 12. Temperature and frequency dependence of the H_8 linewidth of 2'-GMP at pD 7.0 ± 0.2.


OBSERVED LINEWIDTH (Hz)
18
16
14
12
10
8
6
4
2
-30
-20
-10
0
10
20
30
40
50
60
T (°C) →
220 MHz
100 MHz

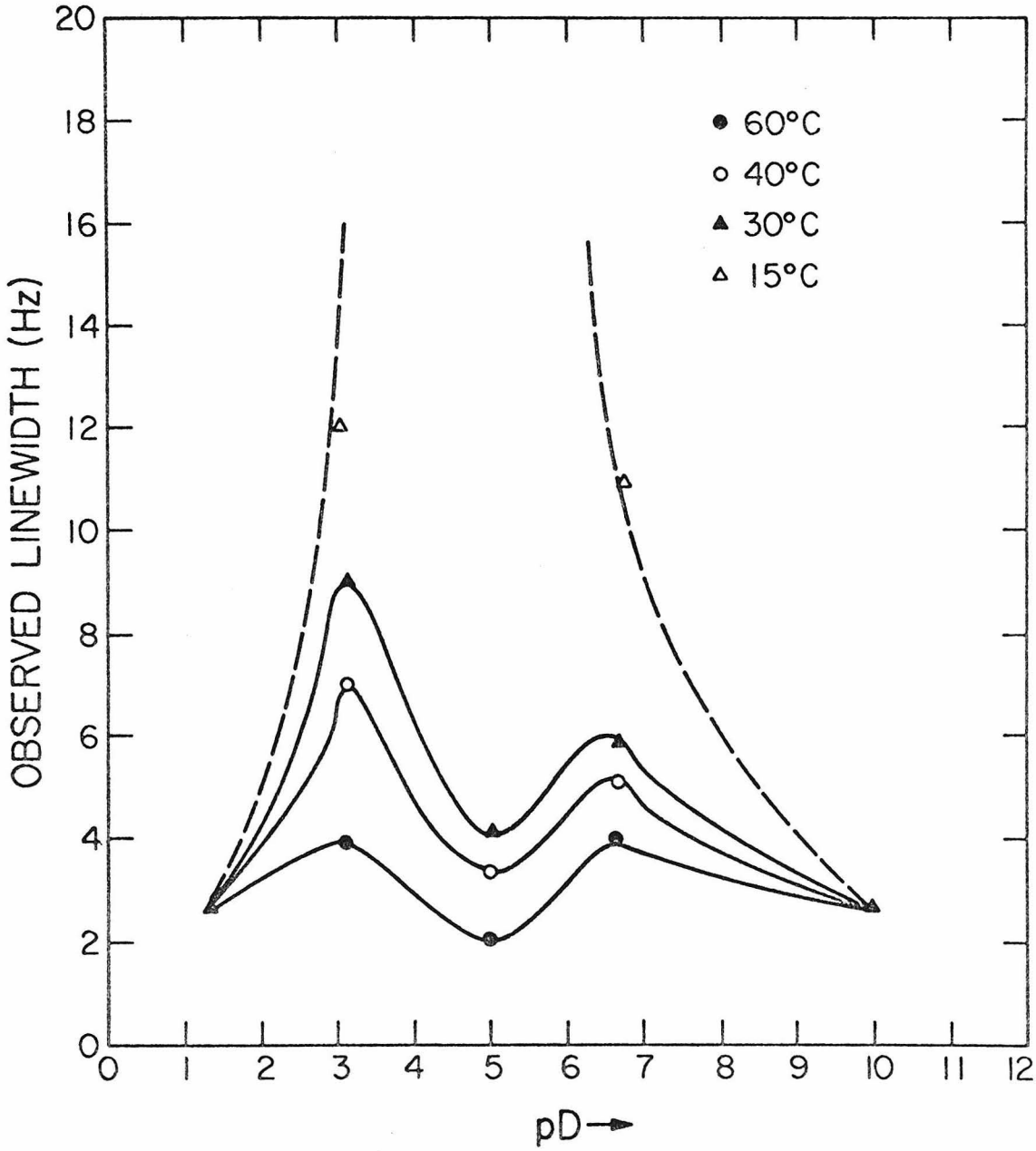
H_8 linewidth becomes essentially independent of field and decreases with decreasing temperature. The greatest line-broadening was observed in the intermediate temperature region. At 220 MHz, the maximum H_8 linewidth is 13.0 - Hz at 10°C and at 100 MHz the H_8 resonance reaches a maximum width of 8.0 Hz at 0°C. Temperature studies of the H_8 linewidth at various solution pD's reveal a similar behavior, except that the temperature at which the observed H_8 line-width is a maximum varies with the pD of the solution. At a given magnetic field, the temperature corresponding to the maximum linewidth was found to be greatest at pD ~ 5.0.

The pD dependence of the H_8 linewidth of a 0.05 M 5'-GMP solution is given at several temperatures in Fig. 13. Perhaps the most striking feature of these data is the abrupt discontinuous increase in H_8 linewidth near 20°C over the pD range 3 to 7. These observations together with the concomitant increase in the viscosity of the sample point to a phase transition from a disordered to a highly ordered system. Outside this pD range and above the "phase transition" temperature, the variation of the H_8 line-width with temperature is similar to that observed for 2'-GMP.

2.3. Intensity Measurements.

Experimentally we have observed that the intensity of the guanine H_8 resonance of 2'-GMP changes with tempera-ture. We have studied this intensity variation in some

Figure 13. Temperature and pD dependence of the H_8 linewidth of 5'-GMP at 0.05 M and 100 MHz.


OBSERVED LINEWIDTH (Hz)
pD→
60°C
40°C
30°C
15°C
0
2
4
6
8
10
12
14
16
18
20
0
1
2
3
4
5
6
7
8
9
10
11
12

detail and have summarized the results in Table III. These intensity measurements were made on a sample containing approximately 0.008 M 2'-GMP and 0.01 M 3'-AMP in a D_2O solution saturated with NaCl at pD 6.4. It was necessary to use 3'-AMP as a secondary intensity standard, as the guanine H_8 resonance of 2'-GMP appears significantly downfield from the ribose H_1' resonance, which could otherwise have been used as the primary intensity standard. In our intensity measurements we have therefore compared the intensity of the guanine H_8 resonance of 2'-GMP with the sum of the adenine H_2 and H_8 resonances of 3'-AMP, and this intensity ratio, $I_{H_8(G)}/[I_{H_2(A)} + I_{H_8(A)}]$, was in turn compared with the intensity ratio of the H_1' resonances between these two nucleotides, that is, $I_{H_1'(G)}/I_{H_1'(A)}$.

At 20°C and higher temperatures, the intensity ratio between the H_8 resonances of 2'-GMP and the average of the H_2 and H_8 resonances of 3'-AMP, namely, $I_{H_8(G)}/\frac{1}{2}[I_{H_8(A)} + I_{H_2(A)}]$, was found to be 0.77 $\pm$ 0.01. This value can be compared with the measured intensity ratio of 0.78 $\pm$ 0.01 for $I_{H_1'(G)}/I_{H_1'(A)}$ at the same temperature. At -10°C, the intensity ratio $I_{H_8(G)}/\frac{1}{2}[I_{H_2(A)} + I_{H_8(A)}]$ was found to be diminished to 0.68 $\pm$ 0.02, whereas the corresponding ratio for the H_1' resonances, $I_{H_1'(G)}/I_{H_1'(A)}$, remained essentially unchanged at 0.76 $\pm$ 0.01. From these measurements it was concluded that the H_8 resonance of 2'-GMP at -10°C accounted for only 88 $\pm$ 3% of the intensity

at 20°C or higher temperatures.

Figure 14. Comparisons of the spectral area between the H_2, H_8, H_1' resonances of 3'-AMP and H_8, H_1' resonances of 2'-GMP at -10°C and 20°C.

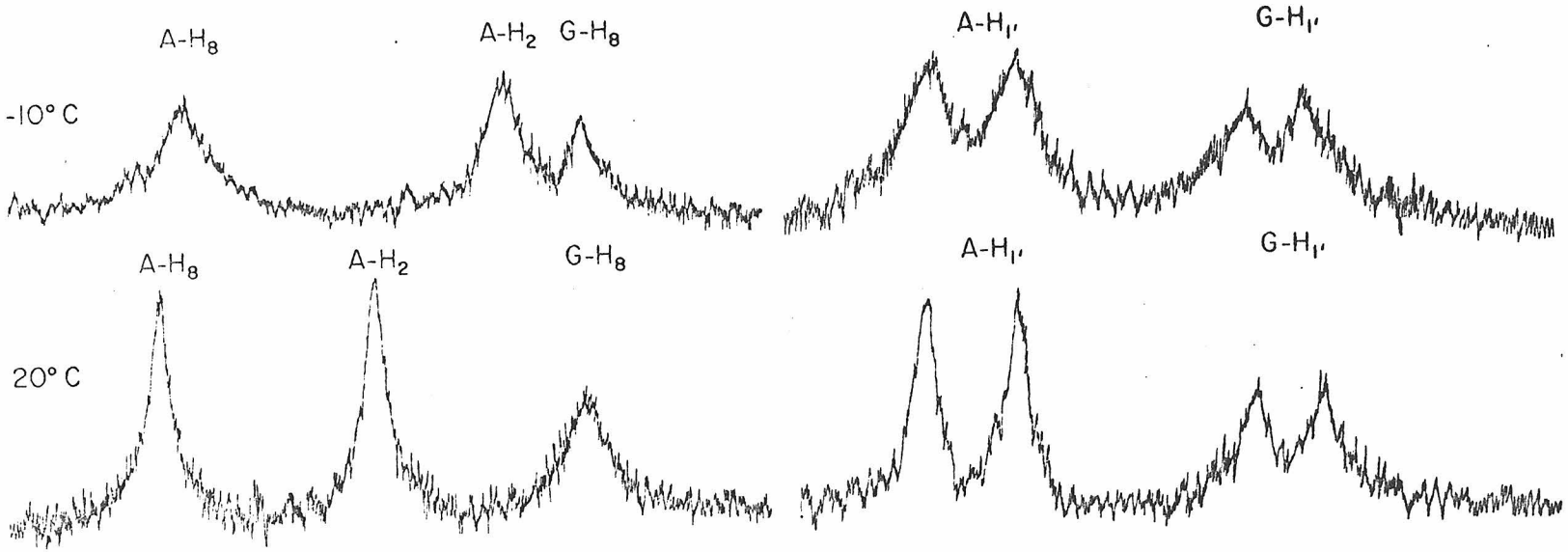

A-H8
A-H2
G-H8
-10° C
A-H1'
G-H1'
A-H8
A-H2
G-H8
20° C
A-H1'
G-H1'

TABLE III

Summary of Intensity Measurements[a]

Temperature	Intensity (Arbitrary Units) H_8(G)	Intensity (Arbitrary Units) H_2(A) + H_8(A)	$I_{H_8}(G) / \frac{1}{2}[I_{H_2}(A) + I_{H_8}(A)]$		Intensity (Arbitrary Units) $H_{1'}$(G)	Intensity (Arbitrary Units) $H_{1'}$(A)	$I_{H_{1'}}(G) / I_{H_{1'}}(A)$	
20°C	56	146	77%		110	140	79%	average
20°C	90	234	77%		150	190	79%	78 ± 1%
20°C	108	276	78%		115	150	77%	
20°C	70	184	76%	average				
20°C	44	110	80%	77 ± 1%				
20°C	55	148	76%					
20°C	115	304	76%					
30°C	60	154	78%					
50°C	55	145	78%					
-10°C	40	120	67%		120	160	75%	average
-10°C	43	130	66%		110	142	77%	76 ± 1%
-10°C	46	130	71%		95	125	76%	
-10°C	38	114	67%	average				
-10°C	20	56	71%	68 ± 2%				
-10°C	29	84	69%					
-10°C	34	102	67%					
-10°C	42	126	67%					
-10°C	69	200	69%					
-10°C	86	268	64%					

[a] Sample: 0.008 M 2′-GMP, 0.01 M 3′-AMP, pD 6.3, saturated NaCl.

V. DISCUSSIONS

1. Cytosine

1.1. Possible Mechanisms for the Linebroadening of the H_5 Resonance of Cytosine Base.

A number of mechanisms could cause broadening of the H_5 resonance of cytosine and its derivatives in aqueous solution. These include contamination by paramagnetic ions, spin coupling of the H_5 proton with neighboring quadrupolar nitrogen-14, extensive cytosine aggregation, and chemical exchange processes with rates observable by nmr.[60-63] The details of these mechanisms have been discussed in chapter II of the part of the thesis. Here we merely present all the experimental evidences which enable us to eliminate the various possibilities and to suggest the most reasonable interpretation of our experimental observations.

(a) Paramagnetic impurities. Coupling between protons and the unpaired electrons of a paramagnetic impurity can significantly broaden the proton resonances. When a paramagnetic ion, such as Co^{+2}, Ni^{+2}, Mn^{+2}, Fe^{+3} is added to a 0.05 M 5'-CMP solution at concentration levels of the order of 10^{-3} to 10^{-4} M, broadening of the H_5 resonance is indeed observed for pD > 4. However, this broadening is not selective for the H_5 resonance, since the H_5 as well as the H_1' resonance is broadened to a

similar extent. Selective broadening of the H_5 resonance has been reported for cytosine derivatives upon the addition of cupric ion at pD 7.0 at concentration levels higher than 2 x 10^{-5} M,[71] and this selective broadening has generally been taken as evidence for Cu^{+2} binding to N_3. However a temperature study of this Cu^{+2} broadening undertaken in conjunction with this work demonstrated that this broadening exhibits quite a different temperature dependence from that observed for an "undoped" 5'-CMP sample. For example, upon cooling the sample to -20°C, the H_5 resonance of a 5'-CMP solution containing 2 x 10^{-5} M of Cu^{+2} remains broad, whereas it has sharpened considerably in the case of an "undoped" 5'-CMP sample. In the Cu^{+2} "doped" solution, the H_5 and H_6 resonances are also broadened to a similar extent at low temperatures. Finally we note that at the concentration level of paramagnetic ions required to observe the Cu^{+2} broadening, EDTA ethylenediamine are 8-hydroxy quinonine are effective in sharpening the H^5 resonance, and hence both chelating agents are effective in the removal of Cu^{+2} and possibly most other paramagnetic ions as well. By contrast, our experiments on the effect of EDTA and ethylenediamine, 8 hydroxyquinonine on the H_5 linewidth in the "undoped" solutions reveal quite a different behavior. Whereas 2 x 10^{-6}M of EDTA will sharpen the H_5 resonance at pD 5.8, a similar effect on the H_5 linewidths is not observed with the

ethylenediamine even at concentrations of 10^{-3} M or greater. On the basis of the above observations, we therefore conclude that paramagnetic ion contamination is not the source of the H_5 broadening observed in the case of the "undoped" cytosine solutions, a result which is, perhaps, not surprising in view of our extensive efforts to eliminate these contaminants in the preparation of our samples. It is our judgement that the purification procedure which we have adopted at least reduces the concentration of these contaminants to a level which does not give rise to paramagnetic broadening.

The possibility of broadening by paramagnetic ions can further be eliminated by the following theoretical considerations. If the broadening arises from the presence of paramagnetic ions, the temperature and frequency (vide-infra) dependences which we have observed would suggest that the rate of chemical exchange of the cytosine base between a complexed and free environment is slow compared with the paramagnetic contact shift at low temperatures, and that at the higher temperatures, the linewidth is controlled by the rate of relaxation through a change in the precession frequency. However, at the concentration levels of paramagnetic ions possible ($\leq 10^{-6}$M), the observed line broadening would imply a nuclear hyperfine coupling constant[72,73] $|A/h| \gg 10$ MHz, a value which is one or two orders of magnitude larger than those which

are normally observed for common paramagnetic ions, a result which we feel is unlikely.

(b) Scalar Coupling to Quadrupolar Nuclei. Indirect spin-spin coupling between the H_5 proton and one or more quadrupolar N^{14}'s can lead to broadening of the H_5 resonance because of quadrupolar relaxation of the nitrogen-14 nucleus. The extent of this broadening depends on the N^{14} coupling constant as well as the relaxation rate of the quadrupolar nucleus. When the quadrupolar relaxation is fast, the H_5 resonance should appear as sharp or broad singlet in the absence of other spin coupling. In the limit of slow N^{14} relaxation, the H_5 resonance should appear as a broad triplet, if only one nitrogen 14 nucleus is involved. On the basis of known N^{15}-H coupling constants[74,75] as well as reasonable estimates of N^{14} relaxation times for this and related molecules, we expect the nitrogen-proton coupling constants to be smaller than the N^{14} relaxation rates in these systems. Hence the former limit should apply, which appears to be the case. However, if this mechanism is the source of the cytosine H_5 broadening, we might expect to observe a dependence of the H_5 linewidths on the viscosity of the solution. The nitrogen relaxation time can be varied by increasing the concentration of the nucleoside or nucleotide in solution (see below) or by increasing the viscosity of the medium through the addition of high concentrations of salt. No dramatic effects were

noted in these experiments. By contrast, the H_5 linewidth exhibits a dependence on the Larmor frequency, an observation which is inconsistent with such an interpretation since there is little question that we are in the "extreme narrowing" limit in these experiments, i.e., $\omega_N{}^{14}\tau_c \ll 1$, and $(\omega_H - \omega_N{}^{14})T_1 \gg 1$.[60] Finally, it would also be difficult to rationalize the striking dependence of the H_5 linewidth on the solution pD as well as the absence of broadening in the case of 4-N, N-dimethylcytosine and 3-methylcytidine, two cytosine derivatives with seemingly rather similar electronic structure to the parent molecule.

The linebroadening of the H_5 resonance of cytosine base, due to the scalar coupling with amino protons or deuterons which undergo chemical exchange with the solvent can be ruled out, in view of the small coupling effect between them.

(c) Molecular Aggregation. Nucleotides and nucleosides, especially those of the purine bases, have been shown to aggregate by vertical stacking in aqueous solution.[76-78] This self-association has been demonstrated for cytosine and its derivatives although it occurs to a somewhat lesser extent. The formation of aggregates results in the increase of the average rotational correlation time of the molecules,[79] which shortens both the longitudinal and transverse relaxation times of the nuclear spins and can lead to a greater observed resonance linewidth if the intrinsic relaxation contribution exceeds the contribution

from field inhomogeneity. For the H_5 and H_6 cytosine protons, the most important contribution to the relaxation rate should be the magnetic dipole-dipole interaction between these protons. Thus, if extensive cytosine aggregation is the origin of the H_5 linewidth, a concomitant linebroadening should also be observed for H_6. In any case, on the basis of the known self-association constant for cytosine derivatives,[76] extensive cytosine aggregation is not expected in the concentration range under consideration. The successive stacking equilibrium constant in aqueous solution is known to be only 0.87 for cytosine. Even at a base concentration of 0.1 M, the average degree of association is only 1.1. The absence of any significant change in the H_5 linewidth over a concentration range from 0.003 M to 0.1 M indicates that either self-association is not extensive or that whatever association does exist does not contribute detectably to the H_5 linewidth.

(d) Chemical Exchange. We now consider chemical exchange as a broadening mechanism. When a molecule can exist in two distinct molecular environments, the chemical shift(s) and the linewidth(s) of the resulting nmr signal(s) will be determined by the ratio of the rate of chemical exchange to the chemical shift difference between the two environments (the nmr time scale). When the exchange is fast compared to the nmr time scale, only the weighted average signal is observed and its linewidth is given by[61]

$$\Delta\nu_{obs} = \Delta\nu_o + 4\pi p_A^2 p_B^2 \Delta^2(\tau_A + \tau_B), \quad (7)$$

where $\Delta\nu_{obs}$ is the observed linewidth; $\Delta\nu_o$ is the weighted average linewidth in the absence of chemical exchange; p_A and p_B are the relative populations of the two exchanging species; τ_A and τ_B denote their pre-exchange lifetimes; and Δ is the chemical shift difference of the proton between the two environments. It is clear that in this limit, increasing the temperature will decrease the linewidth, since both τ_A and τ_B decrease with increasing temperature. In the slow exchange limit, i.e., when the exchange is slow on the nmr time scale, we expect to see two separate signals, provided that the populations of both species are large enough and their resonances narrow enough to enable them to be observed. The linewidth of the resonances due to species A is, for example, given by [61]

$$(\Delta\nu_{obs})_A = (\Delta\nu_o)_A + 1/\pi\,\tau_A \quad (8a)$$

Here $(\Delta\nu_o)_A$ is the resonance linewidth of species A in the absence of chemical exchange. Note that because of the reciprocal dependence on τ_A, the linewidth increases with temperature in this limit. As expected, in the intermediate exchange region, where the exchange is taking place at a rate comparable to the nmr time scale, maximum broadening occurs. Thus, as the exchange rate increases from the slow to the fast exchange limit, the resonance broadens to a

maximum width, and then decreases in width.

The observed variation of the H_5 linewidth of 5'-CMP with temperature (Fig. 8) is characteristic of the spectral behavior depicted above and strongly suggests that the H_5 broadening is due to chemical exchange of the cytosine base between two molecular environments. Presumably, the three temperature regions cited earlier in the results section correspond to the regions of fast, intermediate, and slow chemical exchange. This conclusion is clearly borne out by the effect of frequency and/or magnetic field on the broadening, where the nmr time scale has been altered by varying the magnetic field. Since Δ , the chemical shift difference of the H_5 resonance between the two magnetic environments, is directly proportional to the magnetic field strength in the limit of fast exchange, we expect the exchange linewidth to be proportional to the square of the magnetic field (cf. Eq. (7)). In the limit of slow exchange, however, the linewidth should be independent of the magnetic field (cf. Eq. (8a)). Our experimental data show that in the high temperature region, the linewidth of the H_5 resonance of 5'-CMP increases by a factor of two to four (depending on temperature) as the magnetic field is increased from 32.5 kilogauss (100 M Hz) to 51.7 kilogauss (220 MHz). This agreement with prediction, together with the absence of field dependence in the low temperature region, provides strong evidence in support of broadening

by chemical exchange. The results of our intensity measurements are also consistent with this interpretation. At low temperatures, we observed a significant reduction in the H_5/H_1' intensity ratio, and from this result we concluded that the major species is being monitored and that the minor component is present to the extent of the order of 10%. (Note that in this analysis we are making the assumption that the exchange rate is fast compared with Δ for the H_1' proton in the two environments over the temperature range under consideration.) The H_5 resonance of the minor species is presumably not observed, either because of its relatively low population and/or because the resonance is broad as might be expected. At high temperatures, the observed H_5 to H_1' intensity is unity, as expected for rapid exchange.

A change in the magnetic environment of the H_5 proton can occur in a number of ways, but the fact that the line broadening occurs in cytosine and nucleosides as well as nucleotides indicated that neither the phosphate group nor the sugar moiety is responsible for the chemical shift difference of this proton between the two species. Thus the magnetic nonequivalence must arise from some chemical modification localized in the cytosine base. Since the observed broadening is most pronounced for H_5, indicating that the chemical shift difference between the two molecular environments is greatest for this proton, we surmise that

the structural difference must be confined to that part of the cytosine base which is in close proximity to this proton. The absence of an observed concentration dependence of the H_5 linewidth suggests that the chemical exchange process is unimolecular in cytosine, that the difference in the molecular environment does not arise from the molecular association, and hence the species involved must be tautomeric structures of the neutral cytosine base. An equilibrium between the amino and imino tautomers of cytosine immediately comes to mind. The absence of line broadening of the H_5 resonance in two methylated derivatives, 4-N, N-dimethylcytosine, where the cytosine base can exist only in the amino form, and 3-methylcytidine, where the base is frozen in the imino structure, provides plausibility for this hypothesis. In any case, this latter observation strongly suggests that structural differences at the amino group and the N-3 position of the cytosine base are responsible for the chemical shift difference of the H_5 proton in the two species involved in the chemical exchange. We shall later show that an equilibrium between the amino and imino tautomers of cytosine can indeed account for all our nmr observations, and we shall present a kinetic analysis which bears on the plausibility of this hypothesis.

If a chemical shift nonequivalence exists for the H_5 proton between the two species involved in the chemical exchange, we might expect a chemical shift difference

between the H_6 protons as well, albeit a smaller one. In this connection, we note that a slight broadening is indeed observed for the H_6 resonance. If the two species involved in the chemical exchange correspond to the amino and imino tautomers of cytosine, the chemical shifts of the protons in two species, δ_A and δ_B can be estimated from their chemical shifts in model compounds which exist predominantly in one tautomeric structure. Cytidine is believed to be predominantly in the amino form, so that the observed chemical shifts should approximate δ_A. At pD 8.5, the nonexchangeable methyl group of 3-methylcytidine freezes the cytosine base in the imino form, and the chemical shifts here may be used to approximate δ_B. At 100 MHz, the H_5 and H_6 resonances of 3-methylcytidine are 20 Hz and 4 Hz, respectively, downfield from those of cytidine. Since Δ is 20 Hz for H_5 and only 4 Hz for H_6, the effect of exchange broadening on the linewidth of the H_5 resonance should be much more pronounced than that of H_6, as observed. Thus if our hypothesis regarding the amino-imino equilibrium is correct, we would also have a satisfactory rationale for the greater exchange width observed for H_5 compared to H_6 in normal cytosine.

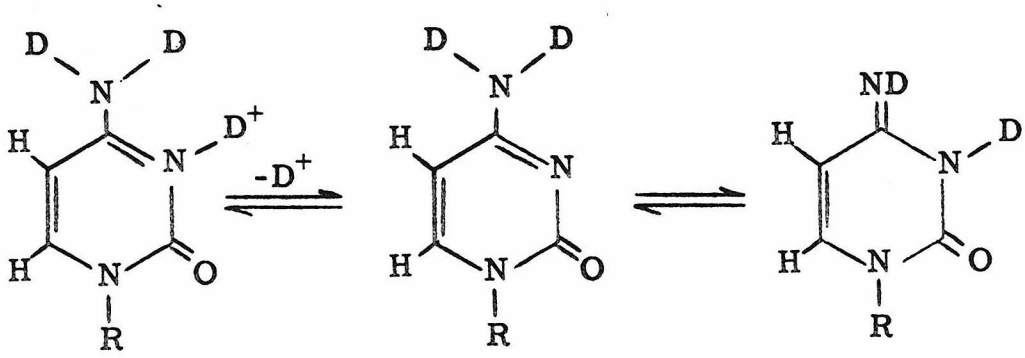
Before concluding this section, it would seem appropriate to comment on the progressive downfield shifts of the H_5 resonance relative to that of H_1' observed with increasing temperature. It is likely that these shifts are also associated with the dynamic equilibrium between the

amino and imino tautomers. If at low temperatures, the tautomeric exchange is slow, then the spectral position of the H_5 resonance should correspond to that of the amino tautomer. As the kinetics becomes more rapid with increasing temperature, we expect gradual collapse of the signals due to the two species (even though in this case, the signal due to the minor tautomer is not apparent) and a shift of the resonances toward a spectral position determined by the weighted average chemical shift, $\delta_{mean} = \delta_A p_A + \delta_B p_B$. It is possible that the $\delta_{H_5} - \delta_{H_1}$, differential shifts observed with increasing temperature merely reflect the influence of these time-dependent factors on the nmr spectrum. At 100 MHz, the H_5 resonance of the imino tautomer is expected to appear ~ 20 Hz downfield from that of the amino tautomer. Our data indicate that the percentage of minor tautomer is ~ 15% at 50°C (vide-infra). Thus, if $\delta_A - \delta_B$ remains unchanged over the temperature range under discussion, we expect a downfield shift of ~ 3 Hz between -20°C and 50°C, if the two temperatures adequately represent the two exchange limits. This calculated shift may be compared with the observed $\delta_{H_5} - \delta_{H_1}$, differential shift of -2.1 ± 0.5 Hz (at 100 Mz) over the same temperature range. The agreement is satisfactory, considering that we have ignored other contributions to $\delta_{H_5} - \delta_{H_1}$, and that the exchange is probably not totally quenched at -20°C. In this connection, we note that almost no differential

$\delta_{H_5} - \delta_{H_1}$, shifts were observed with temperature in 5'-UMP, where tautomerism of the pyrimidine base is known to be absent.

1.2. Kinetic Analysis

To make more plausible our contention that the observed linebroadening of the H_5 resonance arises from chemical exchange between the amino and imino tautomers, we have undertaken a kinetic analysis of this problem. In this analysis, we have considered the following equilibrium



where A, B and C represent the amino and imino tautomers and the protonated cytosine species, respectively. It is necessary to include in this discussion the protonated species in addition to the two neutral tautomers, since our data extend over a wide range of pH. However the chemical exchange between the protonated and the amino tautomer is expected to be fast, so that the problem can still be treated as a two-site exchange problem.

At pD ~ 6.0, the extent of protonation of the

cytosine base is negligible. If we can then assume that the linebroadening arises from the chemical exchange between the two neutral tautomers (A and B), then the linewidth of H_5 resonance can be calculated for a given k_a (the preexchange rate constant for A), if the relative populations between two tautomers as well as the chemical shift difference, Δ , is known. We have carried out such calculations for the H_5 resonance using the Dynamic NMR program developed by G Binsch and D.A. Kleier.(80) Typical results of this calculation are summarized in Figures 15 and 16 . In Figure 15 the calculated linewidth is given as a function of k_a for a range of the chemical shift difference, Δ , assuming 14% minor tautomer. Similarly, in Figure 8 the calculated linewidth is given as a function of k_a for several populations of the minor tautomer in the range 0.08 to 0.20 assuming a Δ of 55 Hz at 220 MHz.

From plots of this type, it is possible to extract the rate constant k_a from the experimental linewidth for a given assumed Δ and ΔG°, where ΔG° is the standard free energy difference between the two tautomers. At a given temperature, the observed linewidths at 100 MHz and 220 MHz must, however, yield the same k_a and this correspondence must apply over the whole temperature range, if Δ and ΔG° are to be unique. (We are assuming here that Δ and ΔG° are invariant over this temperature range.) As a results of a number of iterations, we have found that $\Delta \sim 55$ Hz at 220

Figure 15. Calculated dependence of the H_5 linewidth at 220 MHz on the exchange rate constant k_a (T) for a range of chemical shift difference (Δ) of the H_5 proton between the two tautomers assuming 14% of minor tautomer.

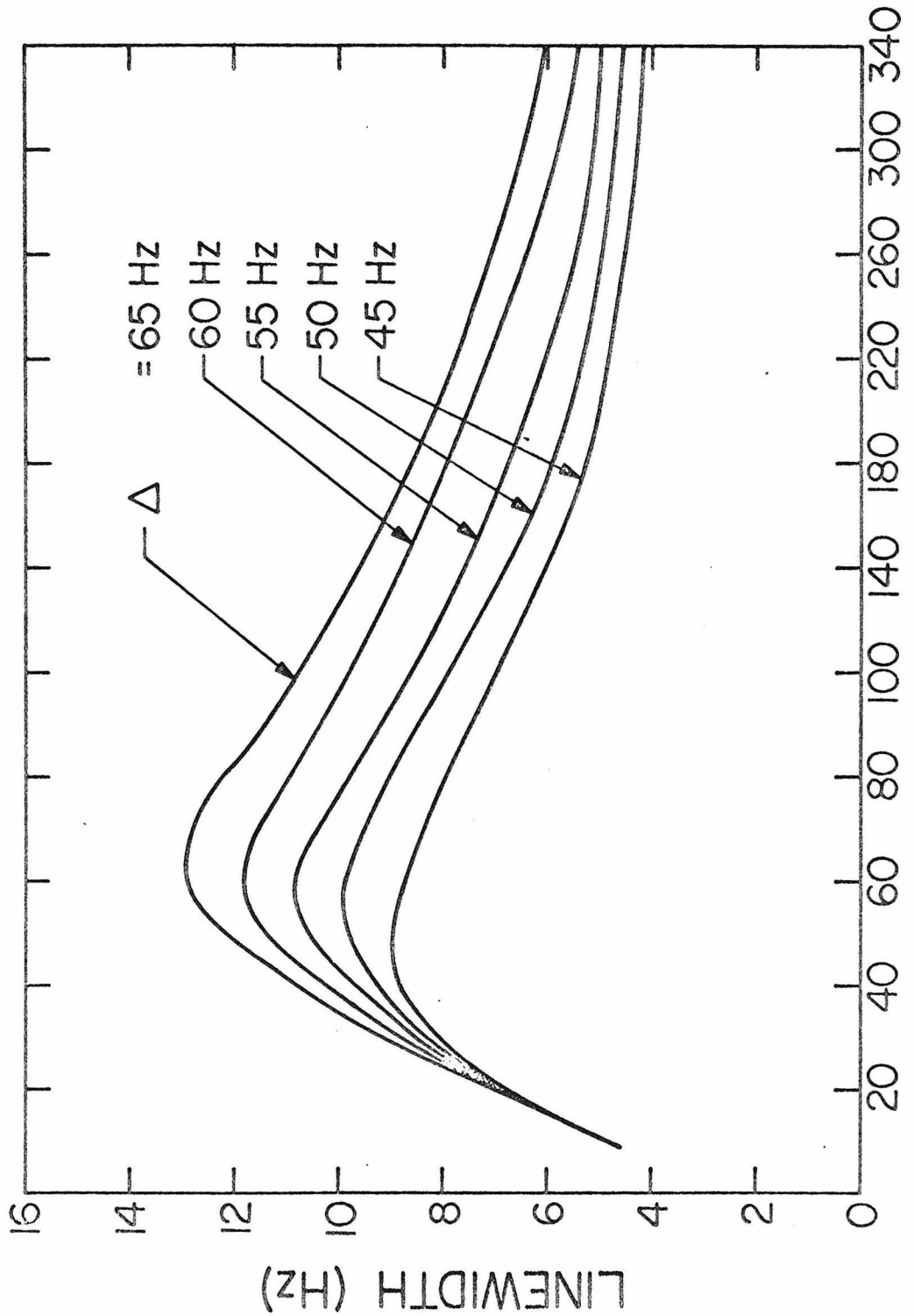
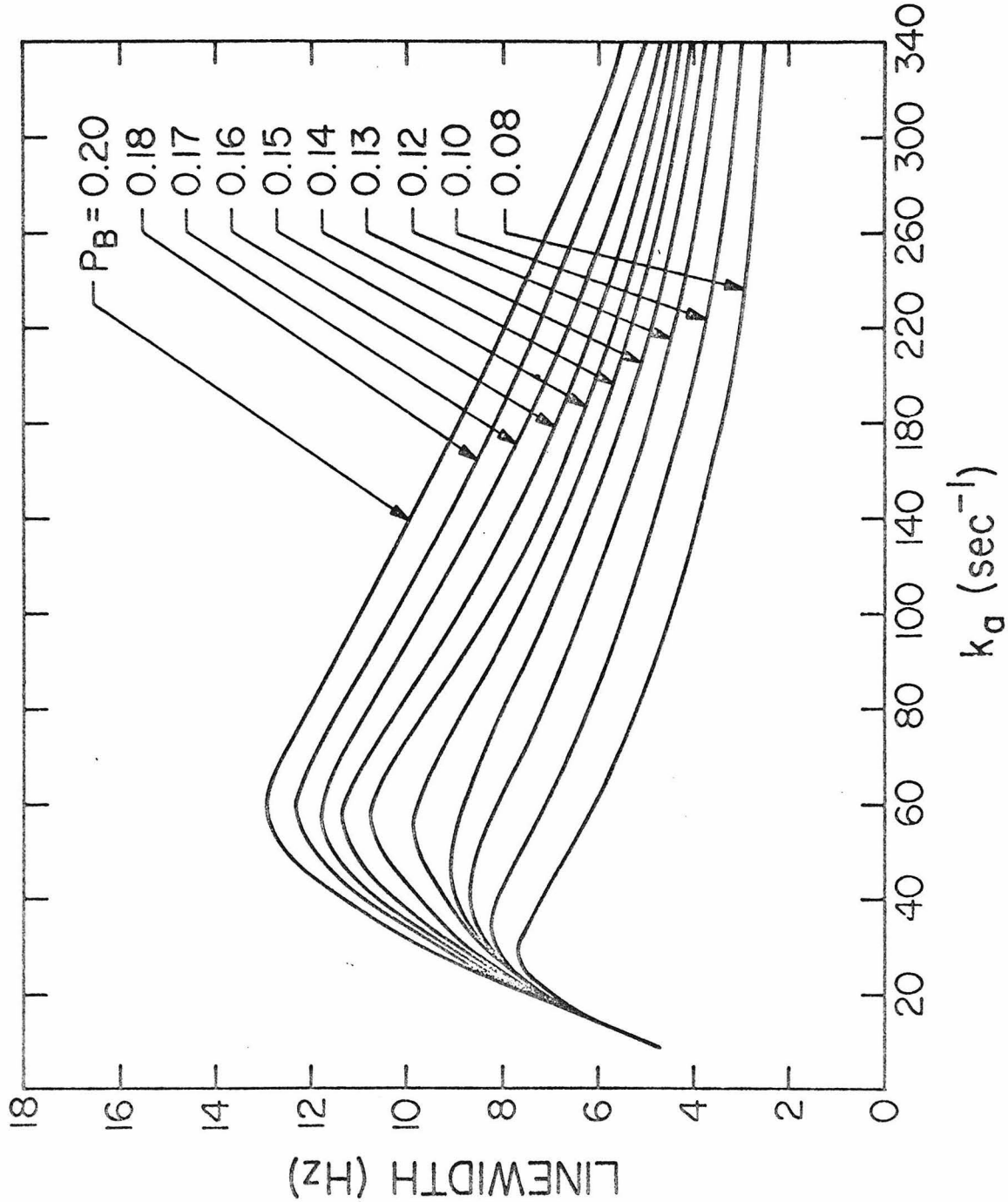

LINEWIDTH (Hz)
0
2
4
6
8
10
12
14
16
20
40
60
80
100
140
180
220
260
300
340
k_a (sec^{-1})
Δ = 65 Hz
60 Hz
55 Hz
50 Hz
45 Hz

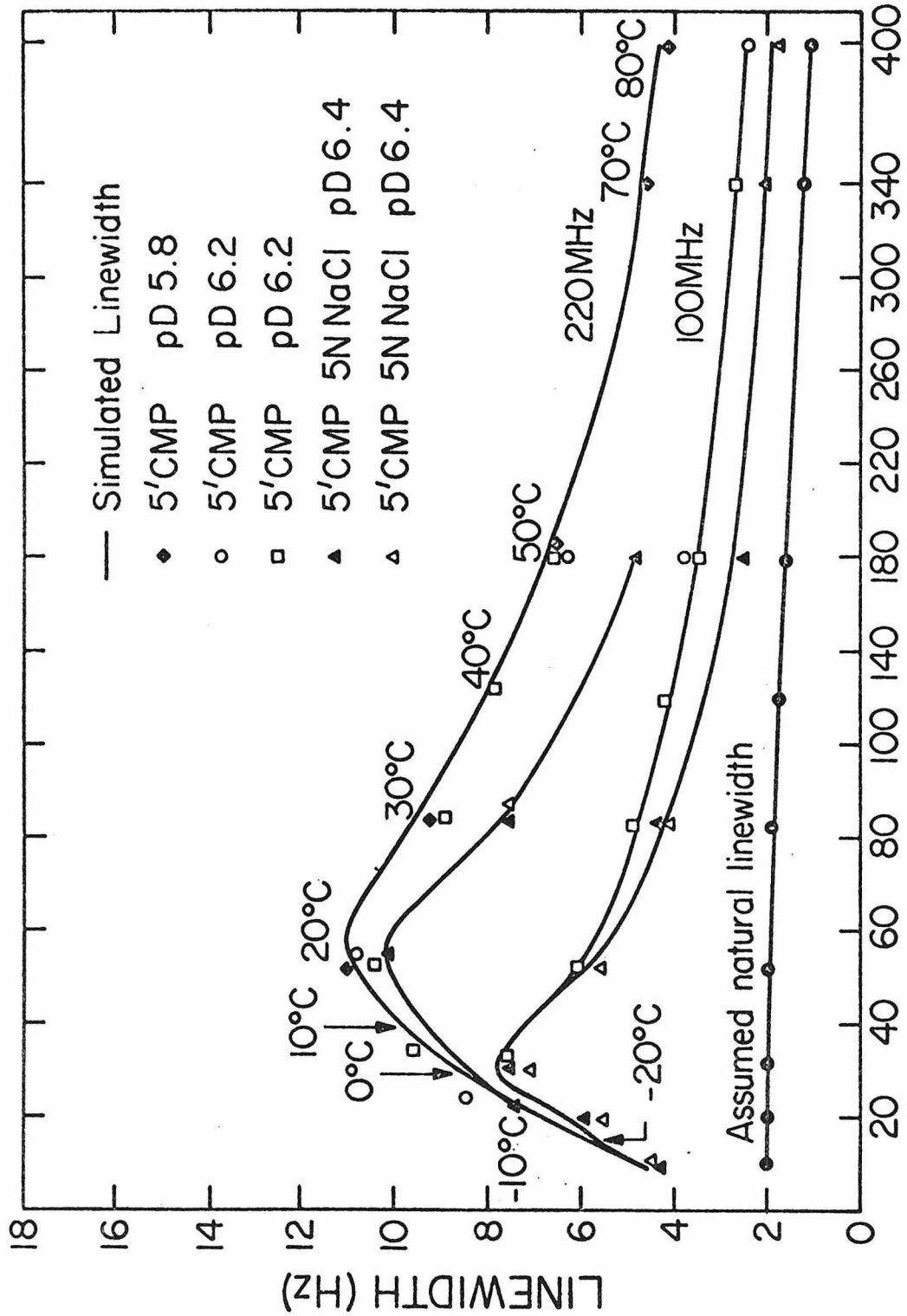
Figure 16. Calculated dependence of the computed H_5 linewidth at 220 MHz on the exchange rate constant k_a (T) for a range of populations of the minor tautomer assuming a Δ of 55 Hz.


LINEWIDTH (Hz)
0
2
4
6
8
10
12
14
16
18
20
40
60
80
100
140
180
220
260
300
340
k_a (sec^{-1})
P_B = 0.20
0.18
0.17
0.16
0.15
0.14
0.13
0.12
0.10
0.08

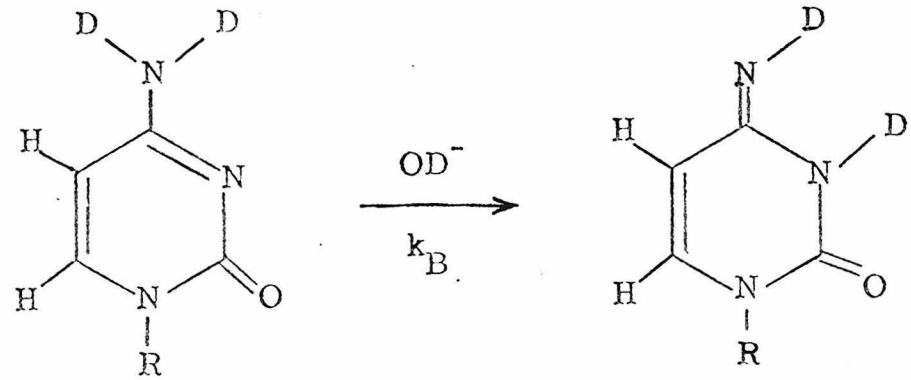
MHz and $\Delta G^{\circ} \backsim 1.1$ Kcal/mole gave the best fit to our data. A comparison of the observed linewidths at the two frequencies with the theoretically expected linewidth is depicted in Fig. 17 for 5'-CMP near pD $\backsim$ 6.0 and for 5'-CMP in 5 N NaCl at pD $\backsim$ 6.4. We recall that our intensity measurements of the H_5 resonance had previously suggested that the percentage of the minor tautomer is 10 $\pm$ 2% at -20°C, which corresponds to a $\Delta G^{\circ} \backsim 1.1$ kcal/mole, and studies of model compounds had suggested that the chemical shift difference between the two tautomers, Δ, is $\backsim$ 0.20 ppm. These values agree well with our kinetic analysis and indicate that to a good approximation $\Delta S^{\circ} \backsimeq 0$ and $\Delta H^{\circ} \approx 1.1$ kcal/mole.

In Figure 18 , we have plotted the log of the rate constant, $\log_{10} k_a$, so extracted versus the reciprocal absolute temperature for 5'-CMP at pD $\backsim$ 6 both in the presence and absence of NaCl. If the chemical exchange between the two tautomers involves only a single mechanism log k_a vs 1/T should yield a straight line with slope given by $\Delta E^{\ddagger}/R$, where $\Delta E^{\ddagger}$ is the activation energy for the exchange process. Experimentally, we have found that this plot is not linear throughout the whole temperature range; instead, k_a appears to be the sum of two exponentials. This suggests the possibility that two activation energies are involved and the exchange between the two tautomers might involve two parallel paths. A careful study of the relative contributions of the two terms indicates that the

Figure 17. A comparison of the calculated and observed H_5 linewidths at 100 MHz and 220 MHz for 5'-CMP in aqueous solutions near pD ~ 6.0 and for 5'-CMP in 5 N NaCl at pD ~ 6.4.

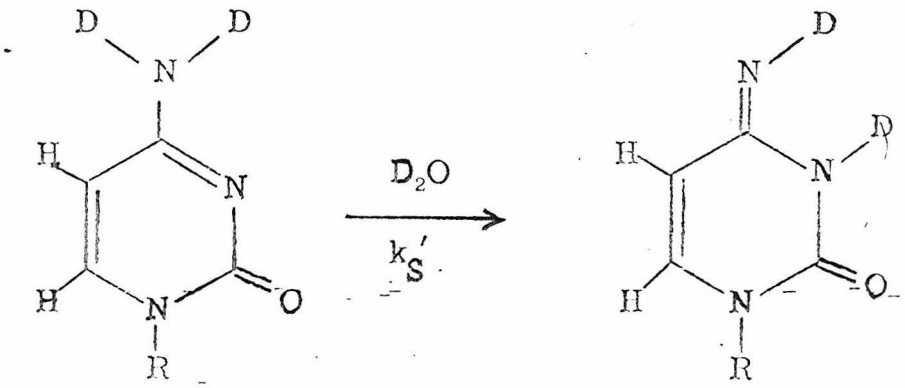

Simulated Linewidth
5'CMP pD 5.8
5'CMP pD 6.2
5'CMP pD 6.2
5'CMP 5N NaCl pD 6.4
5'CMP 5N NaCl pD 6.4
LINEWIDTH (Hz)
k_a (sec^{-1})
220MHz
100MHz
Assumed natural linewidth
10°C
20°C
0°C
30°C
40°C
50°C
70°C
80°C
-10°C
-20°C

higher activation energy component is pH independent, but solvent dependent, e.g., the salt concentration, whereas the lower activation energy component is pD dependent only. In view of these observations, and the pronounced pD dependence of the H_5 linewidth, particularly at higher pD's, it is reasonable to propose that the chemical exchange in one of these paths is base (OD^-) catalyzed, i.e.,



A B

and that other kinetic pathway involves the solvent, i.e.,



If this kinetic model is correct, then

$$-\left(\frac{d[A]}{dt}\right)_f = k_B\,[A]\,[OD^-] + k'_S\,[A]\,[D_2O]$$

$$= k_B\,[A]\,[OD^-] + k_S\,[A] \quad , \qquad (9)$$

$$\text{and } k_a(T) = \frac{1}{\tau_{ab}} = -\left(\frac{1}{[A]}\,\frac{d[A]}{dt}\right)_f$$

$$= k_B\,[OD^-] + k_S$$

$$= A_B\,[OD^-]\,e^{-\Delta E_B^{\neq}/RT} + A_S\,e^{-\Delta E_S^{\neq}/RT} \quad , \qquad (10)$$

where $-\left(\frac{d\,A}{dt}\right)_f$ is the pre-exchange rate of the major tautomer, k_B and k_S denote the rate constants for the base and solvent catalyzed steps, respectively; $k_S = K'_S\,[D_2O]$, and $\Delta E_B^{\neq}$ and $\Delta E_S^{\neq}$ denote the activation energies for the base and solvent catalyzed steps. Analysis of the data in Fig. 18. yielded the following activation energies:

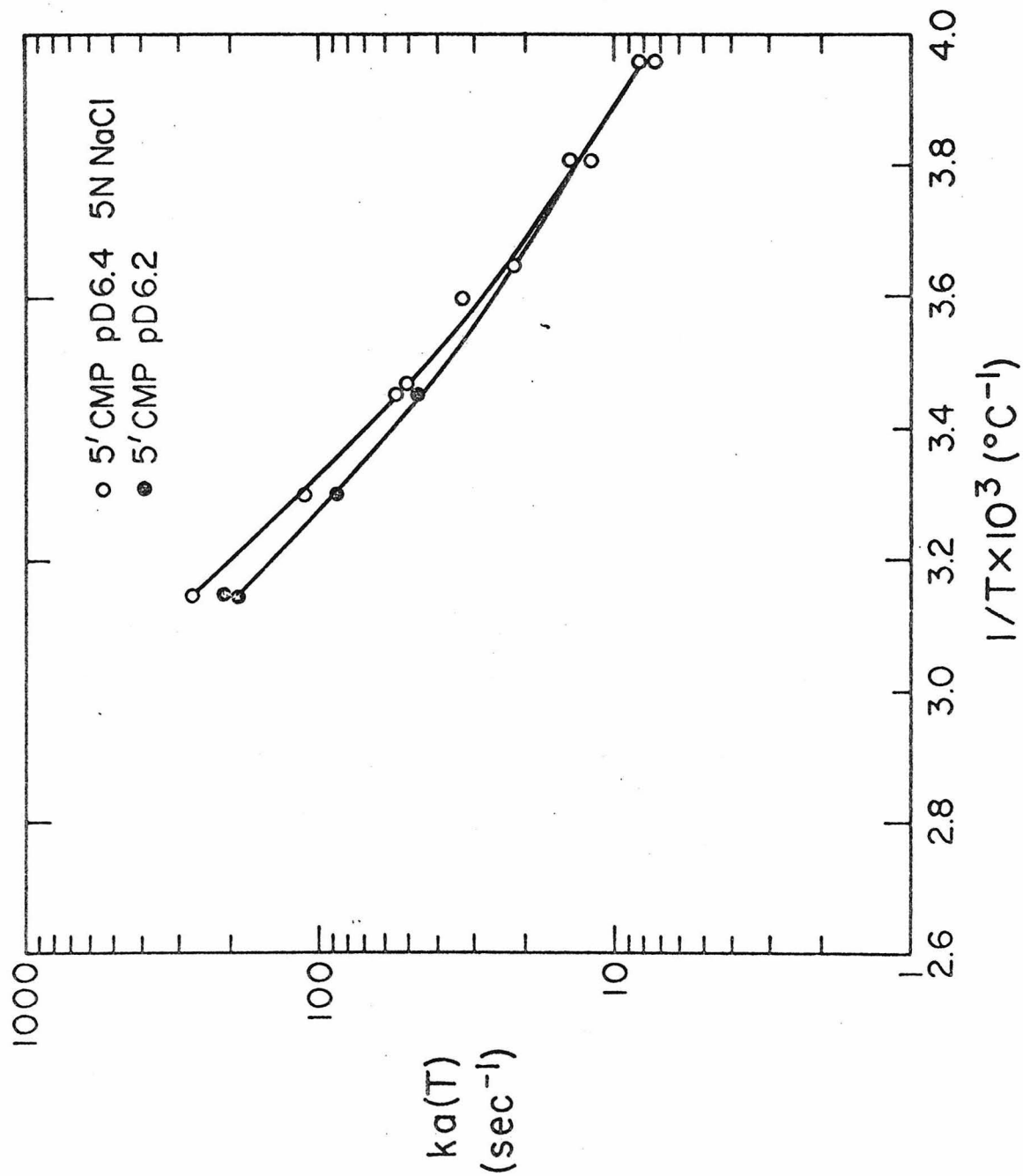
$\Delta E_B^{\neq} = 3.5 \pm 0.5$ kcal/mole.

$\Delta E_S^{\neq} = 7.8 \pm 0.6$ kcal/mole (in the absence of salt).

$\Delta E_S^{\neq} = 10.8 \pm 0.5$ kcal/mole (in the presence of 5N NaCl).

We have examined the base catalyzed step of this tautomeric exchange in detail. The pD dependence of the linewidth comes about in two ways. First, the exchange rate $k_a(T)$ involves the base (OD^-) concentration directly.

Figure 18. $\log_{10} k_a$ versus 1/T.


o 5'CMP pD6.4 5N NaCl
• 5'CMP pD6.2
ka(T) (sec⁻¹)
1000
100
10
1
2.6
2.8
3.0
3.2
3.4
3.6
3.8
4.0
1/T×10³ (°C⁻¹)

However at lower pD's, where the cytosine base can become protonated, the population of the neutral tautomers p_A and p_B varies with the solution pD. p_A and p_B can be expressed in terms of the solution pD and the protonation constant K_a. Since

$$p_A + p_B + p_C = 1 \quad , \tag{11}$$

$$K_a = \frac{p_A [D^+]}{p_C} \quad , \tag{12}$$

and

$$p_A = p_B \, e^{\Delta H^\circ / RT} \quad , \tag{13}$$

$$p_B = \left\{ 1 + (1 + \frac{[D^+]}{K_a}) \, e^{\Delta H^\circ / RT} \right\}^{-1} \quad . \tag{14}$$

The effective time-scale of the nmr experiment, i.e., Δ_{eff}, is also pH dependent as a result of the rapid chemical exchange between the major tautomer A and the protonated cytosine species C and the effect of this exchange on the spectral position of the weighted average signal due to these species. Δ_{eff} is readily shown to be

$$\Delta_{eff} = \delta_B - \delta_{AC}$$

$$= \delta_B - (p_A \delta_A + p_C \delta_C)/(p_A + p_C) \quad . \tag{15}$$

A knowledge of p_A p_C, p_B, Δ_{eff} and the rate expression k_a (T) enables us to simulate the pD dependence at a given temperature, when the relative contribution of the two

mechanisms is known, i.e., if we know the ratio of the pre-exponential factors, A_B/A_S, or the ratio of the rate constants k_B/k_S at a given temperature.

We have determined the ratio of the pre-exponential factors for the two exchange mechanisms, $\eta = A_B/A_S$, from the pD value corresponding to the observed maximum linewidth at 30°C. To do this, we have used the following approximation for the linewidth,

$$\Delta\nu_{obs} = \Delta\nu_0 + \frac{4\pi p_B^2}{p_A} (1 - p_B)^2 \Delta^2_{eff} \tau_{ab} \quad , \qquad (16)$$

which we have derived from the Bloch equation for three site exchange, i.e., C ⇌ A ⇌ B, in the limit of very fast exchange between C and A and moderately fast exchange between A and B, a condition which is satisfied at 30°C, pD 6.0 and at 100 MHz. We have shown that under these conditions this equation predicts a linewidth which is in good agreement (within 10%) with that obtained by the density matrix method employed in the Dynamic NMR program.(80) Substituting Eqs. (12)-(15) into (16), we obtain the following pD dependence for the H_5 resonance at a given temperature T under conditions where Eq. (16) is applicable and $p_B \ll p_A$.

$$\Delta\nu_{obs} = \Delta\nu_0 + \frac{4\pi}{A_S}\,\Delta_{eff}^2\; e^{(\Delta E_S^{\neq} - 2\Delta H)/RT}$$

$$\times \frac{[D^+]}{\{1 + [D^+]/K_a\}\{[D^+] + \eta\, K_w\, e^{(\Delta E_S^{\neq} - \Delta E_B^{\neq})/RT}\}} \qquad (17)$$

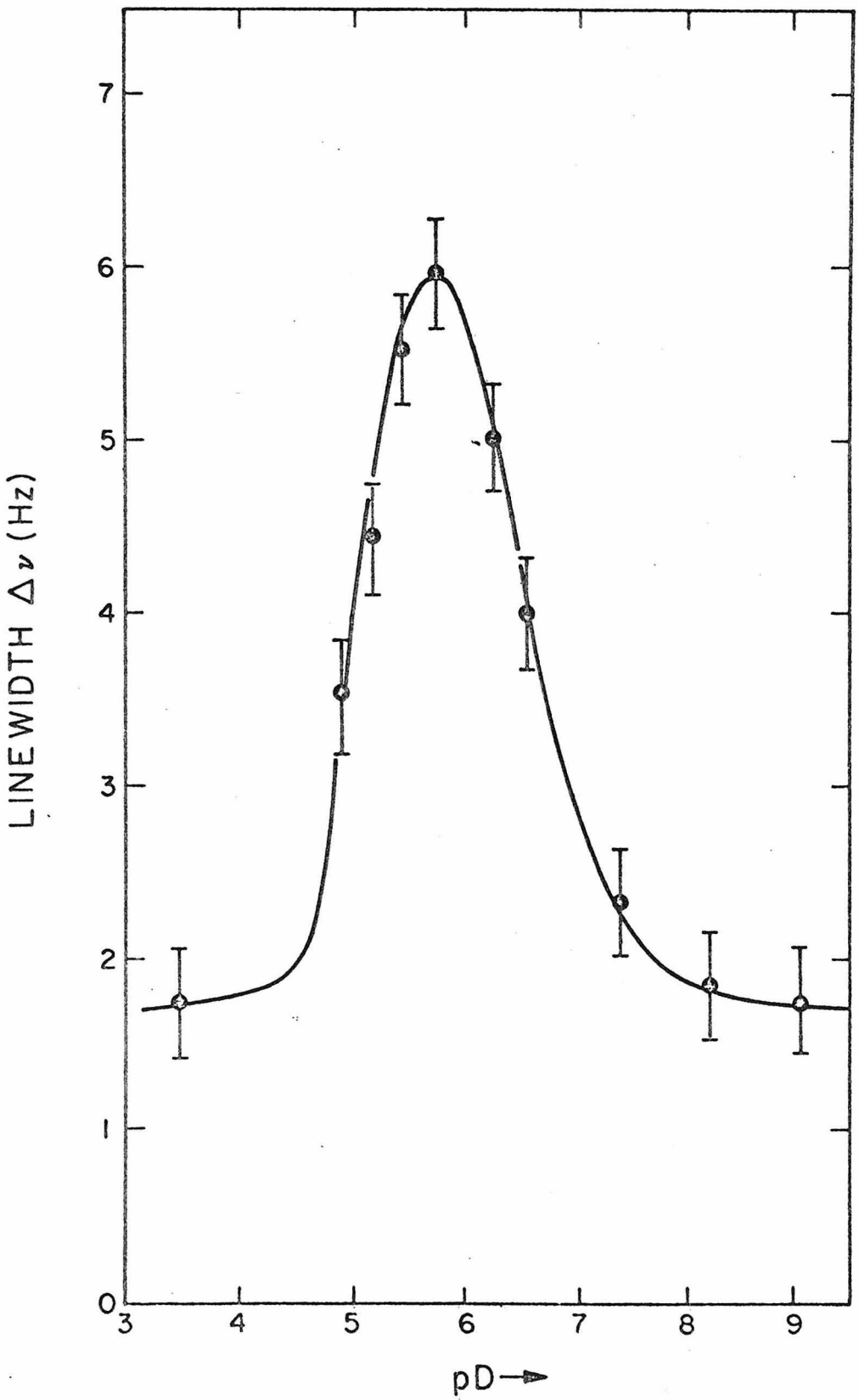
K_w is the ionization constant of water. The pD corresponding to the maximum linewidth can be obtained by differentiating the above expression with respect to $[D^+]$ and setting the resultant derivative to zero. Experimentally, the maximum H_5 linewidth was observed at pD $\sim$ 6.0 at 30°C, from which we ascertain that

$$\eta = 2 \times 10^8\; e^{(\Delta E_B^{\neq} - \Delta E_S^{\neq})/0.6} \qquad (18)$$

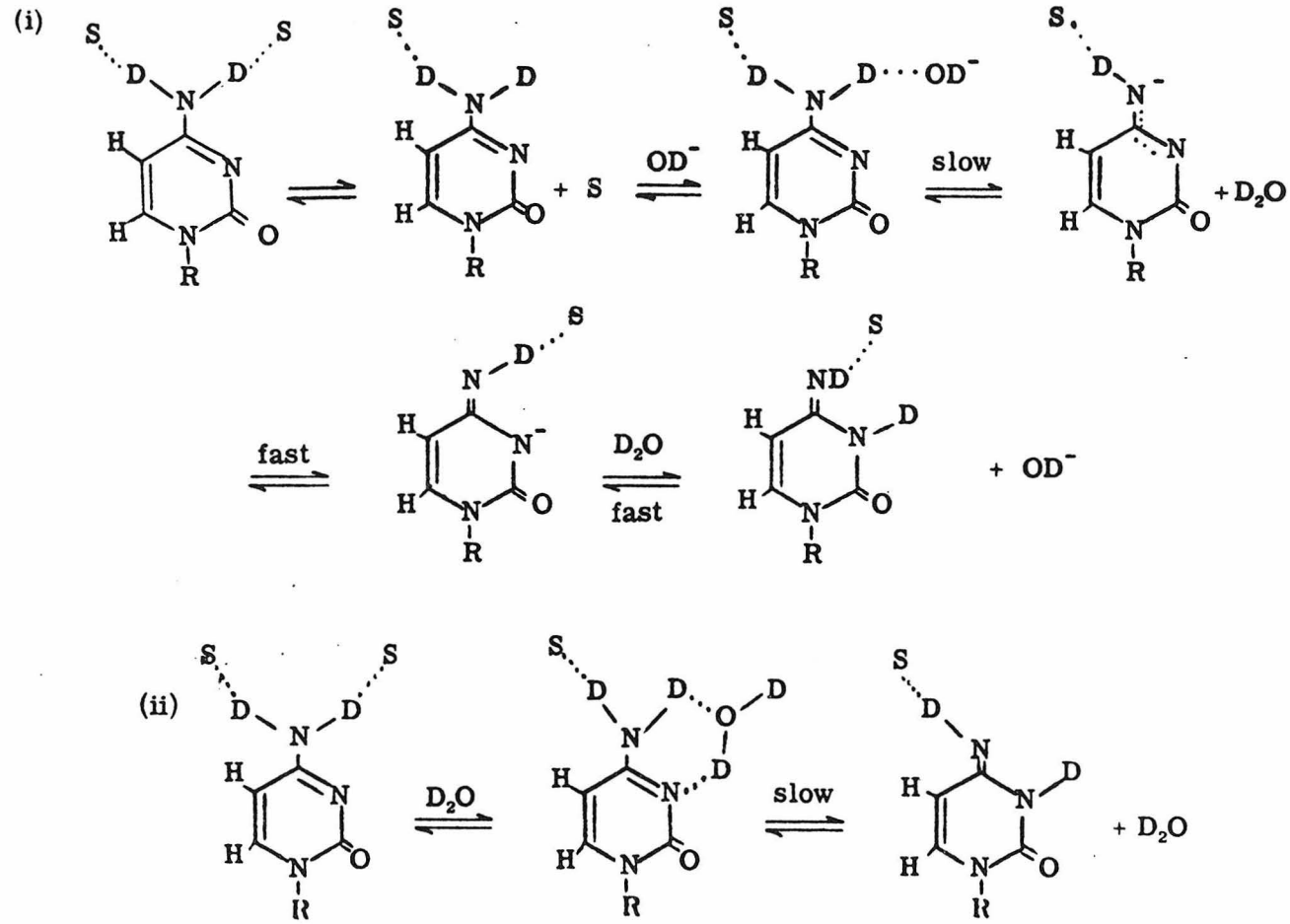
The calculated pD dependence of the H_5 linewidth at 30°C in D_2O solution is depicted in Fig. 19. The overall agreement between the theoretical curve and the experimentally observed linewidths can be seen to be excellent.

The results of this kinetic analysis yielded the following rate constants for the tautomeric exchange process at 30°C. For the base catalyzed step $k_B \sim 1 \times 10^{10}$ ℓ mole^{-1} sec^{-1}; for the step involving the solvent $k'_S \sim$ 1.2 ℓ mole^{-1} sec^{-1}. The rate law suggests the following

Figure 19. A comparison of the calculated and observed pD dependence of the H_5 linewidth of 5'-CMP in D_2O (— theoretical linewidth, I observed linewidth).


LINEWIDTH Δν (Hz)
0
1
2
3
4
5
6
7
3
4
5
6
7
8
9
pD→

mechanism :



Both k_B and k_S' are of the proper order of magnitude[81] for proton abstraction from an amino group by OD^- and D_2O. For example, the rate constant for the base catalyzed step is approximately an order of magnitude lower than that measured for proton abstraction from the methylammonium ion by hydroxyl ion, whereas the solvent catalyzed step is two orders of magnitude faster than the deprotonation of methylammonium ion by H_2O. This difference in the protolysis rate constants between the methylammonium ion and the cytosine base may reflect the acidity of the amino protons in the cytosine base as a result of the partial double bond character between C_4 and N_4. The extra acidity of these protons will result in their stronger interaction with the solvent, accounting for the greater ease with which they are removed by the solvent as well as the less favorable proton abstraction by hydroxyl ion.

2. Guanine

2.1. Mechanisms for the Linebroadening of the H_8 Resonance of Guanine Base.

A number of mechanisms could cause broadening of the H_8 resonance of guanosine and its derivatives. These include contamination by paramagnetic impurities, spin coupling of the H_8 proton with neighboring N^{14} nuclei, extensive molecular aggregation, and chemical exchange processes with rates observable by nmr. The details of

these mechanisms have been discussed in Chapter II of this part of the thesis.

In our present work, we can rule out paramagnetic ion complexation as a source of linebroadening in view of our extensive efforts to remove such possible contaminants in the preparation of our samples. The fact that we do not observe broadening of the H_8 resonance in some of the guanine derivatives as well as several other related compounds, e.g., inosine, 5'-IMP, etc. would seem to rule out the presence of significant concentrations of paramagnetic impurities in the solvent system used in these studies, since the same D_2O was used throughout all the sample preparations. Moreover, with the exception of EDTA, the effects of which we shall comment on later, the addition of efficient chelating agents, such as ethylenediamine, and 8-hydroxy-quinoline, even at concentration levels of 10^{-3} M, does not lead to a reduction of the H_8 linewidth. The linewidth of the H_8 resonance of 2'GMP was also studied in H_2O instead of D_2O solution at pH 5.6. The observation of the similar H_8 linebroadening of 2'GMP seems to rule out the possibility that the linebroadening arises from the paramagnetic contaminants in D_2O solution. Effects arising from scalar coupling of the H8 proton to the two quadrupolar N^{14} nuclei of the imidazole ring of the guanine base can also be eliminated, as there is no similar broadening of the H_8 resonances in inosine,

1-methyl-gaunosine and 6-methoxy-purine riboside, all purine derivatives with seemingly rather similar electronic structure insofar as the imidazole ring is concerned.

These considerations leave molecular aggregation and chemical exchange processes as the only possible source of the H_8 linebroadening. In the case of 2'-GMP, 2-N, N-dimethylamino-methylene-guanosine and ApG, linebroadening by molecular aggregation can further be ruled out since the H_8 linewidth is independent of the concentration. In the case of the other guanine derivatives, with the exception of 1-methyl-guanosine and 6-methoxy-purine riboside where no H_8 linebroadening was observed, molecular aggregation does affect the H_8 linewidth, but our data here indicate that this intermolecular association leads to a sharpening rather than a broadening of the H_8 resonance at temperatures above the phase transition for the gel formation. In fact, under these conditions, the H_8 linewidth of guanosine, 5'-GMP, etc., approaches that observed for 2'-GMP as the solution is diluted. For all the guanosine derivatives, except 1-methyl-guanosine and 6-methoxy-purine riboside, therefore, there is an intrinsic linebroadening of the H_8 resonance at infinite dilution which can only be accounted for by chemical exchange.

Since the H_8 broadening in 2'-GMP is concentration dependent, we first discuss this relatively simple case. Here, the observed variation of the H_8 linewidth with

temperature depicted in Fig.12 leaves little doubt that the linebroadening is due to chemical exchange of the guanine base between two chemically distinct environments. The three temperature regions cited earlier in the results section correspond to the regions of fast, intermediate and slow chemical exchange. This conclusion is clearly borne out by the effect of frequency and/or magnetic field on the linebroadening, where the nmr time scale has been altered by varying the magnetic field. As expected, a strong field dependence of the H_8 linewidth was observed in the high temperature region, where the exchange rate is fast compared to the nmr time scale. At the low temperature region, where the exchange rate is slow, the observed H_8 linewidth is expected to be field independent, as observed. In this limit, one might expect to see distinct resonances for the individual species involved in the chemical exchange. But if the percentage of the minor species is low, the resonances for the minor species would be much broader than those of the major species and might escape detection because of low signal intensities. However, if the percentage of the minor species is not too low, we might expect some reduction in the integrated intensity of the resonances of the major species at low temperatures from that at higher temperatures, to an extent commensurate with the population of the minor species. We do observe such an intensity reduction. The

results of our intensity measurements indicate a reduction of some 12 ± 3% in the intensity of the H_8 resonance of 2'-GMP at -10°C compared to that at 20°C or higher temperature.

A change in the magnetic environment of the H_8 proton can occur in a number of ways, but the fact that the same linebroadening occurs at low concentration in the nucleosides as well as the nucleotides suggests that the sugar-phosphate moiety is not responsible for the chemical shift difference of this proton between the two species. The magnetic nonequivalence, therefore, most likely arises from some modification localized in the guanine base. In view of our recent findings regarding the tautomerism of the cytosine base, a similar tautomerism of the guanine base immediately comes to mind. For guanine nucleosides and nucleotides there are two possible tautomeric exchange processes which might occur: (i) a lactam-lactim exchange involving proton transfer from N_1 to the C_6 keto oxygen; and (ii) an amino-imino exchange involving proton transfer between the amino group and N_3. However, since the H_8 broadening is still observed in the case of 2-N, N-dimethylamino-methylene guanosine, where the amino-imino tautomerism is quenched, but is absent in the case of 1-methyl-guanosine, where the guanine base is frozen in the lactam structure, tautomeric exchange between the lactam and lactim forms of the guanine base is the most probable

source of the observed guanine H_8 line broadening. The absence of H_8 broadening in 6-methoxy-purine riboside, where the base is frozen in the lactim structure, supports this conclusion. A comparison of the H_8 chemical shift in guanosine with that in 1-methyl-gaunosine as well as the earlier ir work of Miles indicates that the major tautomer corresponds to the lactam form.

The pD dependence of the H_8 linewidth of 2'-GMP can be explained as follows. At high pH's, the chemical exchange is expected to be base catalyzed, as the presence of base will aid the removal of a proton from an N-H group or OH group of the guanine base. If this is the case, both the forward and backward rates will be accelerated with increasing pD and correspondingly the H_8 linewidth will be reduced. At low pD's, the protonation of the guanine base occurs for both tautomeric species around pD 2 to 3. (pD titration of model compounds suggests similar pK_a's for both structures.) It should be noted, however, that this protonation process does not occur at the sites involved in the lactam-lactim exchange, since the sites of protonation on the guanine base are generally thought to be N_3 or N_7. Nevertheless, the H_8 linewidth should decrease precipitously below pD $\sim$ 3.0, since we expect the tautomeric exchange to be grossly accelerated when the guanine base becomes positively charged. Consideration of the H_8 chemical shifts for the two tautomers and their respective protonated species

also suggests that the effective chemical shift difference between the lactam and lactim structures decreases with lowering pD. This would lengthen the nmr time scale of observation and contribute to decreasing the observed H_8 linewidth. Finally, it is interesting to note the "plateau" in the H_8 linewidth versus pD profile over the pD range 4 to 6. This behavior can be contrasted with the rather abrupt pD dependence over the same pD region in the case of cytosine. We believe that this difference merely reflects the higher pK_a of the cytosine base compared to the guanine base. We shall later argue that the flat maximum in the 2'-GMP pD profile corresponds to the solvent contribution to the kinetics of the tautomeric exchange.

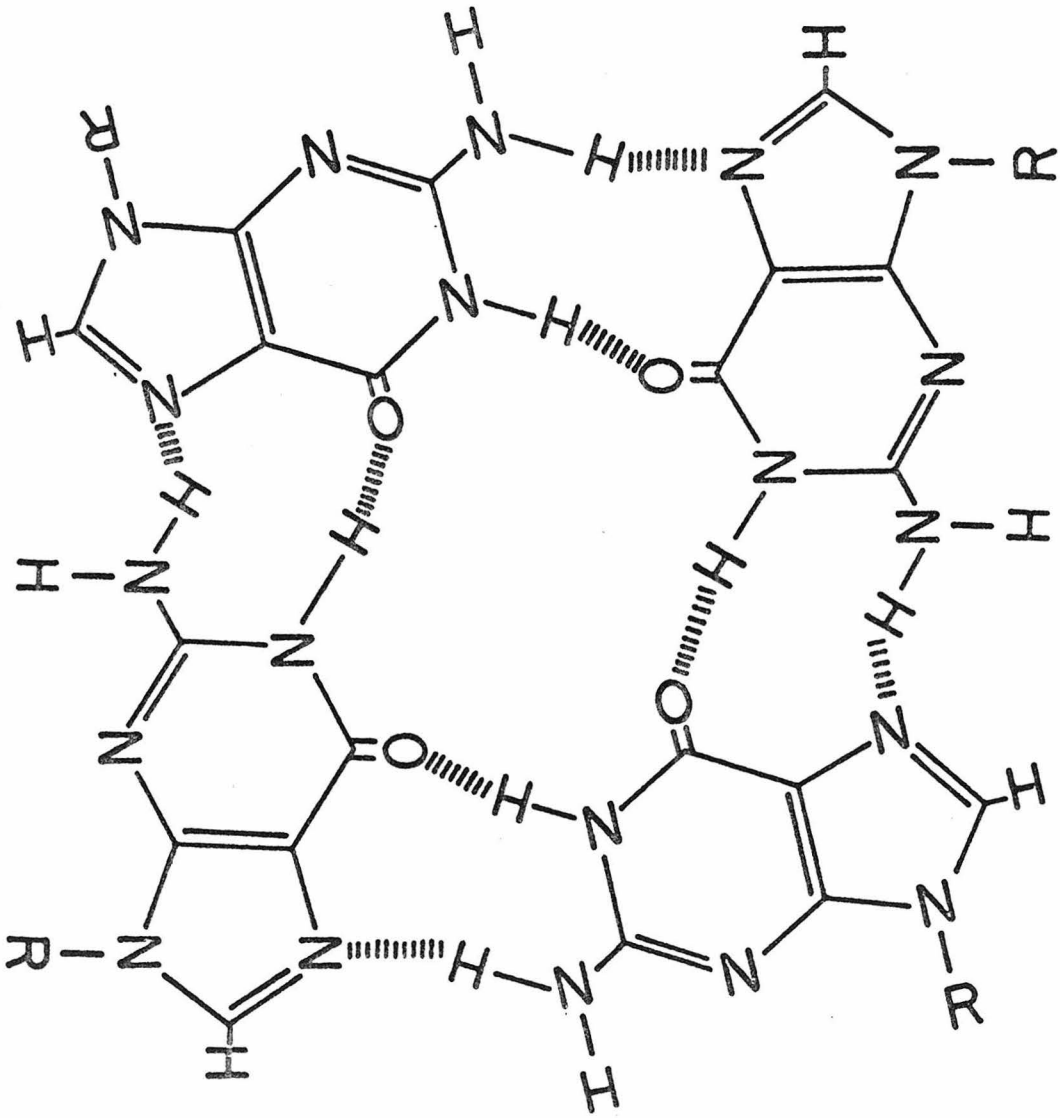
Some discussion of the possible origin of the chemical shift difference of the H_8 proton between the two tautomeric structures would be appropriate at this point. Although the subject of substituent effects on proton chemical shifts in heterocyclic systems is a complicated one and it would be rather difficult to predict accurately the effect of the lactam-lactim tautomerism on the H_8 chemical shift, there is at least one contribution to this shift which may be ascertained with some reliability. It has been argued that since the pyrimidine base is not aromatic when the guanine base is in the keto structure, the induced motions of the π-electrons of the pyrimidine ring in a magnetic field are rather localized so that the

ring current magnetic anisotropy of the guanine base arises principally from the ring current of the π-electron cloud of the imidazole ring. Evidence for this is apparent from the unusual high-field position of the H_8 resonance in the pmr spectrum of guanosine compared with such completely conjugated systems as adenine nucleosides and nucleotides. However, upon transfer of a proton from N_1 to the C_6 keto oxygen to form the lactim structure, the pyrimidine ring becomes conjugated and its π-electrons should contribute additional ring current magnetic anisotropy to the guanine base and lead to a downfield shift of the H_8 proton. Estimation of this ring current effect[61] as well as consideration of the H_8 chemical shifts in model compounds where the guanine base can exist in one or the other tautomeric structure suggests that this additional ring current effect can amount to $\sim$ 30 Hz at 100 MHz.

We now turn to the more complicated case of 5'-GMP. Here the observed pD dependence of the H_8 linewidth at infinite dilution can also be attributed to tautomeric exchange. At higher concentrations, however, and particularly in the pD region 3 to 6, it is clear that some intermolecular process is affecting the tautomeric equilibrium and hence the H_8 linewidth. In view of the strong tendency of 5'-GMP towards gel formation, this molecular aggregation is to be expected, even though our experiments are undertaken at temperatures above the phase

transition of the gel and at concentrations significantly lower than those where gel formation is supposed to prevail. Presumably this intermolecular association is also taking place in solutions of guanosine, 3'-GMP, 2'-deoxyguanosine, 5'-dGMP, and 5'-GTP, albeit to a lesser extent, judging from the effect of concentration on the H_8 linewidths in these guanine derivatives. It has been proposed, at least in the case of 3'-GMP, that the gel formation proceeds via hydrogen-bonding among four guanine bases to form a tetramer (Fig. 20), and that these tetramers in turn stack to form higher aggregates. In the case of 5'-GMP, x-ray studies of fibers obtained from dehydration of the gels has also suggested the possibility of an alternate structure,[69] wherein one of the hydrogen-bonds between the guanine bases in the tetramer is broken to form similar hydrogen-bonds between contiguous tetramers in the "stack" followed by some tilting of the guanine bases to yield a more or less continuous helix. Although the structure of the molecular aggregates for 5'-GMP may indeed differ somewhat from that postulated for 3'-GMP in the gel or fiber state, it is not clear whether these structural differences can also be extended to their structures in solution. Moreover, if the guanine species in equilibrium in solution are principally those of the monomer, the tetramer, and aggregates of the tetramer, it is unnecessary to concern ourselves with these structural differences, as we have no way of ascertaining

Figure 20. Proposed structure of the hydrogen-bonded etramer of guanine derivatives (see Ref. 69 of text).



them in the thermodynamic considerations which we shall present later. The work of Gellert et al.[69] has suggested that the tetramer and its aggregated species are only stable in the intermediate pD region 3 to 6, where the phosphate group of a nucleotide would be protonated ($pD < 6.5$) and when the guanine base is neutral ($3.0 < pD < 9.0$). Should this be the case, and if in addition the lactam tautomeric structure predominates* in the tetramer and its higher molecular aggregates, the intermolecular aggregation would shift the equilibrium away from the lactim species. In solution, we expect the equilibrium between the monomer, the tetramer and its aggregated species to be rapid, i.e., the time scale of breakup of the aggregates should be fast compared with their nmr transverse relaxation times. If this condition is satisfied, then the molecular association should lead to a narrowing of the H_8 resonance, as observed.

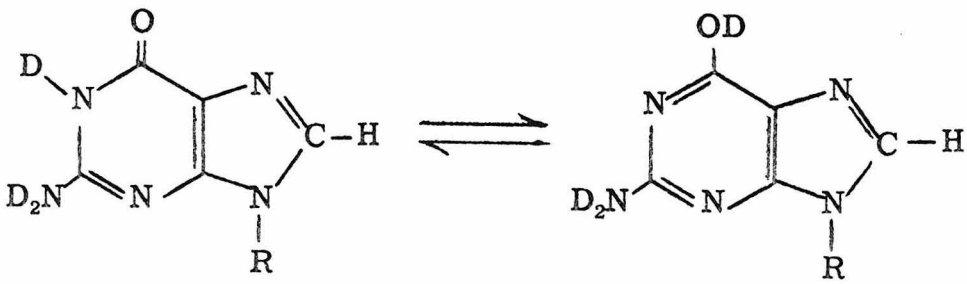
2.2. Kinetic analysis

In order to satisfy our contention that the observed linebroadening of the H_8 resonance of guanosine and its related derivatives arises from tautomeric exchange between the lactam and lactim tautomers of the guanine base, we have carried out a detailed kinetic analysis of this

*Although it is possible to write down a tetramer structure with each of the four guanine bases in the lactim form, this structure can be shown to be much less stable from thermodynamic arguments.

problem similar to that which we have undertaken for the cytosine system.

We consider the relatively simple case of 2'-GMP, where molecular aggregation has been shown to be unimportant and therefore may be neglected, and where the H_8 line-broadening would be predominantly determined by the chemical exchange between the lactam and lactim tautomeric species. Except at low pD's where the guanine base becomes protonated, our observation should then be adequately interpreted in terms of the following equilibrium:



(A) (B)

Assuming this equilibrium, the linewidth of the H_8 resonance can be calculated for a given k_a (the pre-exchange rate constant for A), if the relative populations between the two tautomers as well as the chemical shift difference, Δ, is known. We have carried out such calculations for the H_8 resonance using the Dynamic NMR program developed by G. Binsch and D.A. Kleier.[80] Typical results of this

calculation are summarized in Figs. 21 and 22. In Fig. 21 the calculated linewidth is given as a function of k_a for a range of the chemical shift difference Δ assuming 16% minor lactim tautomer. Similarly, in Fig. 22 the H_8 linewidth is calculated as a function of k_a for several populations of the minor tautomer, p_B, ranging from 0.12 to 0.18 assuming a Δ of 0.30 ppm.

From plots of this type, we may extract the rate constant from the experimental H_8 linewidths for a given assumed Δ and ΔG°, where ΔG° is the standard free energy difference between the two tautomers. At a given temperature, the observed linewidths at 100 MHz and 220 MHz must, however, yield the same k_a and this correspondence must apply over the whole temperature range, if Δ and ΔG° are to be unique. (We are assuming here that Δ and ΔG° are invariant over this temperature range.) As a result of a number of iterations, we have found that $\Delta \backsim 66$ Hz at 220 MHz and $\Delta G^\circ \backsim 1.0$ kcal/mole gave the best fit to our data. A comparison of the observed H_8 linewidth at both 100 MHz and 220 MHz with the theoretically predicted linewidth is depicted in Fig. 23 for 2'-GMP at pD 7.0 $\pm$ 0.2 in saturated NaCl. We recall that our intensity measurements of the H_8 resonance had previously suggested that the percentage of the minor tautomer is 12 $\pm$ 3% at -10°C, which corresponds to a ΔG° of $\backsim 1.0 \pm 0.2$ kcal/mole, and we had suggested that $\Delta \backsim 0.3$ ppm. These values agree

Figure 21. Calculated dependence of the H_8 linewidth at 220 MHz on the exchange rate constant k_a (T) for a range of chemical shift difference (Δ) of the H_8 proton between the two tautomers assuming 16% minor tautomer.

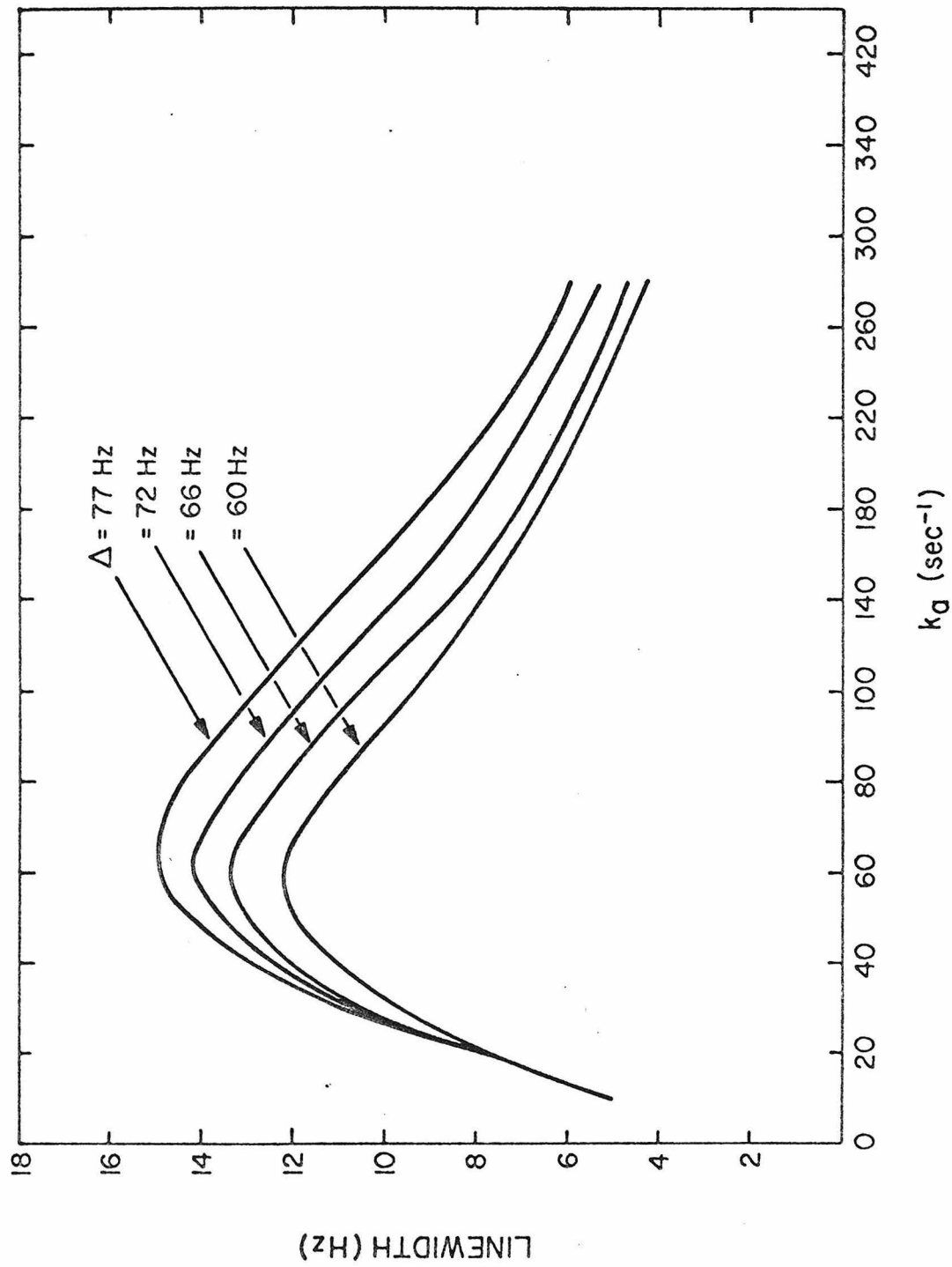

Δ = 77 Hz
= 72 Hz
= 66 Hz
= 60 Hz
LINEWIDTH (Hz)
18
16
14
12
10
8
6
4
2
0
20
40
60
80
100
140
180
220
260
300
340
420
k_a (sec^{-1})

Figure 22. Calculated dependence of the H_8 linewidth on the exchange rate constant k_a (T) for a range of populations of the minor tautomer assuming a Δ of 0.30 ppm.

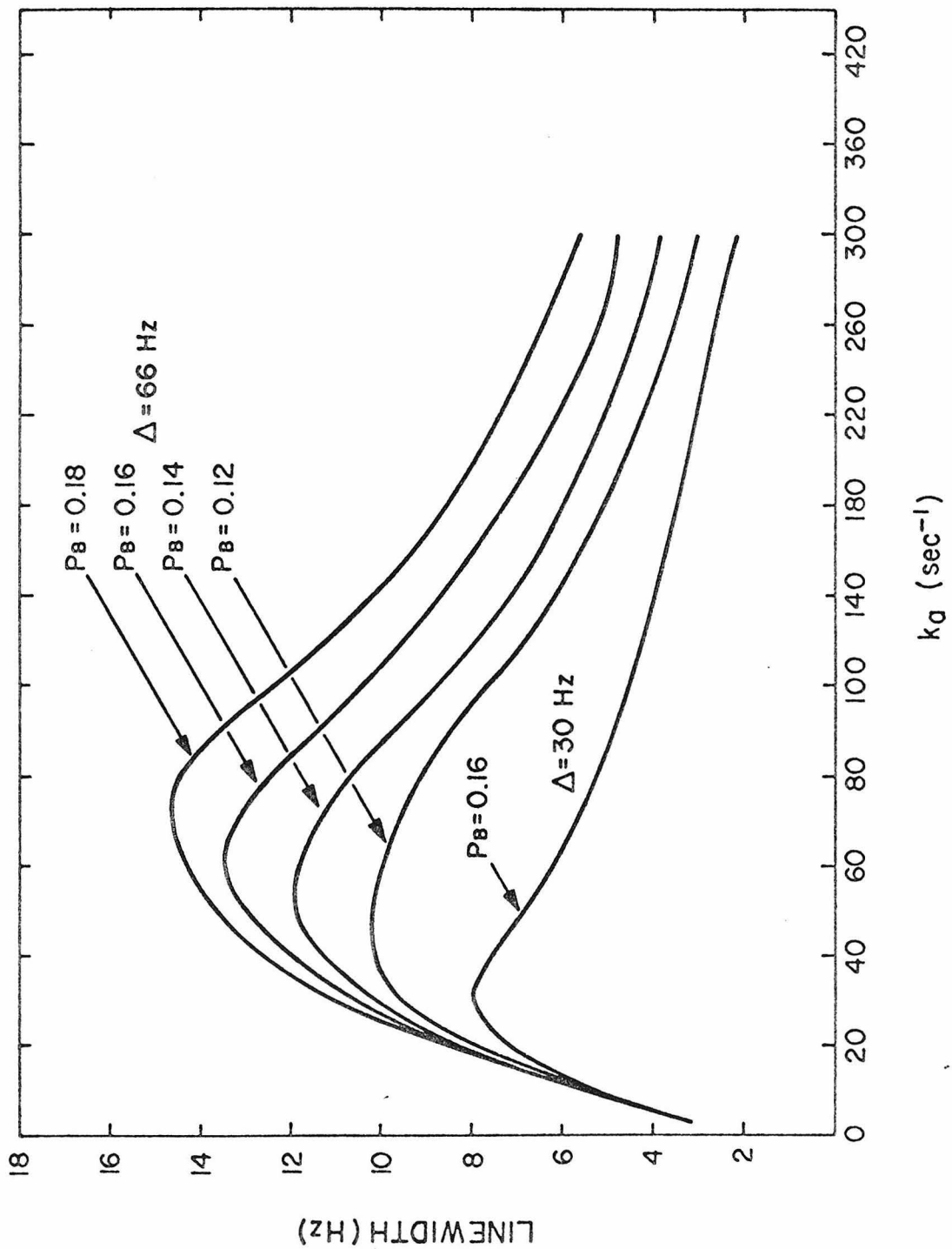
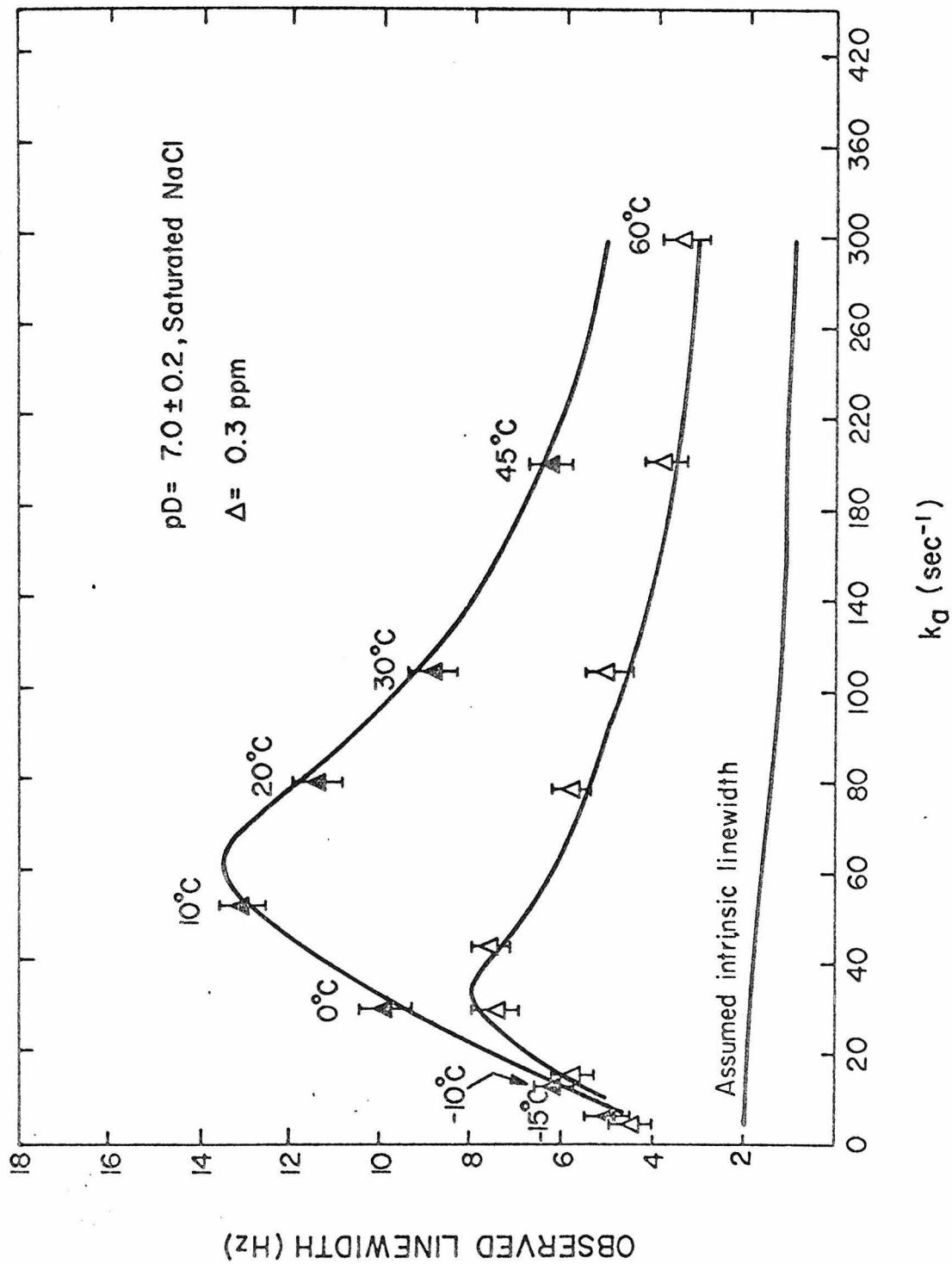

LINEWIDTH (Hz)
18
16
14
12
10
8
6
4
2
0
20
40
60
80
100
140
180
220
260
300
360
420
k_a (sec^{-1})
P_B = 0.18
P_B = 0.16
P_B = 0.14
P_B = 0.12
Δ = 66 Hz
P_B = 0.16
Δ = 30 Hz

Figure 23. A comparison of the calculated and observed H_8 linewidth at 100 MHz and 220 MHz for 2'-GMP near pD $\sim$ 7.0 in saturated NaCl.


pD= 7.0 ± 0.2, Saturated NaCl
Δ= 0.3 ppm
10°C
20°C
30°C
45°C
60°C
0°C
-10°C
-15°C
Assumed intrinsic linewidth
OBSERVED LINEWIDTH (Hz)
k_a (sec^{-1})
18
16
14
12
10
8
6
4
2
0
20
40
60
80
100
140
180
220
260
300
360
420

well with the results of our kinetic analysis and also indicate that to a good approximation $\Delta S^{\circ} \approx 0$ $\Delta H^{\circ} \approx \Delta G^{\circ} \approx 1.0$ kcal/mole.

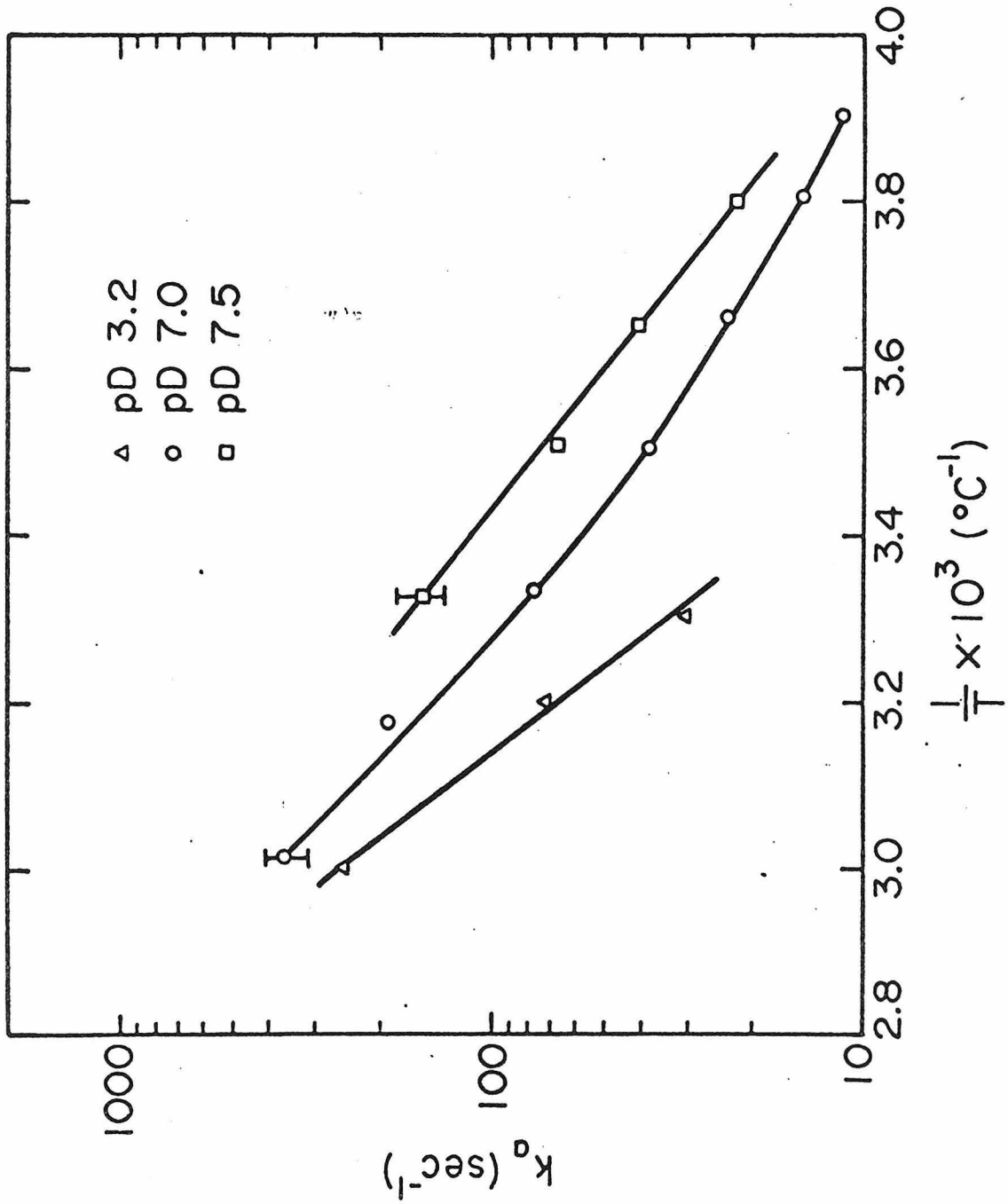
In Fig. 24, we have plotted the log of the rate constant, $\log_{10} k_a$, extracted from the above analysis against the reciprocal of the absolute temperature for 2'-GMP at pD 7.0 ± 0.2 in the presence of saturated NaCl. For comparison, we have included similar data at pD = 3.2 and pD = 7.5. Two exchange paths are indicated by the nonlinearity of the plot for pD = 7.0 as well as the linearity of the plots at pD = 3.2 and pD = 7.5. In view of the strong pD dependence of the H_8 linewidth at high pD's (pD > 6), and the lack of dependence of the linewidth on the solution pD over the pD range 4 to 6, we suggest that the tautomeric exchange under consideration is catalyzed by both the solvent (D_2O) and base (OD^-); i.e.,

$$k_a(T)^{-} = \frac{1}{\tau_{ab}} = -\left(\frac{1}{[A]} \frac{d[A]}{dt}\right)_f$$

$$= k_B [OD^-] + k_S' [D_2O] = k_B [OD^-] + k_S \quad , \qquad (19)$$

and that the pD dependence of the $\log_{10} k_a$, vs 1/T plots merely reflects the relative contribution of the two mechanisms at various pD's. At pD = 7.5, base catalysis

Figure 24. $\log_{10} k_a$ versus the inverse of the absolute temperature for 2'-GMP.


pD 3.2
pD 7.0
pD 7.5
1000
100
10
2.8
3.0
3.2
3.4
3.6
3.8
4.0
k_o (sec^{-1})
$\frac{1}{T} \times 10^3$ (°C^{-1})

should dominate, whereas this mechanism should be insignificant at pD = 3.2 for any reasonable value of k_B. Under these conditions, $\log_{10} k_a$ versus 1/T should be linear, as observed.

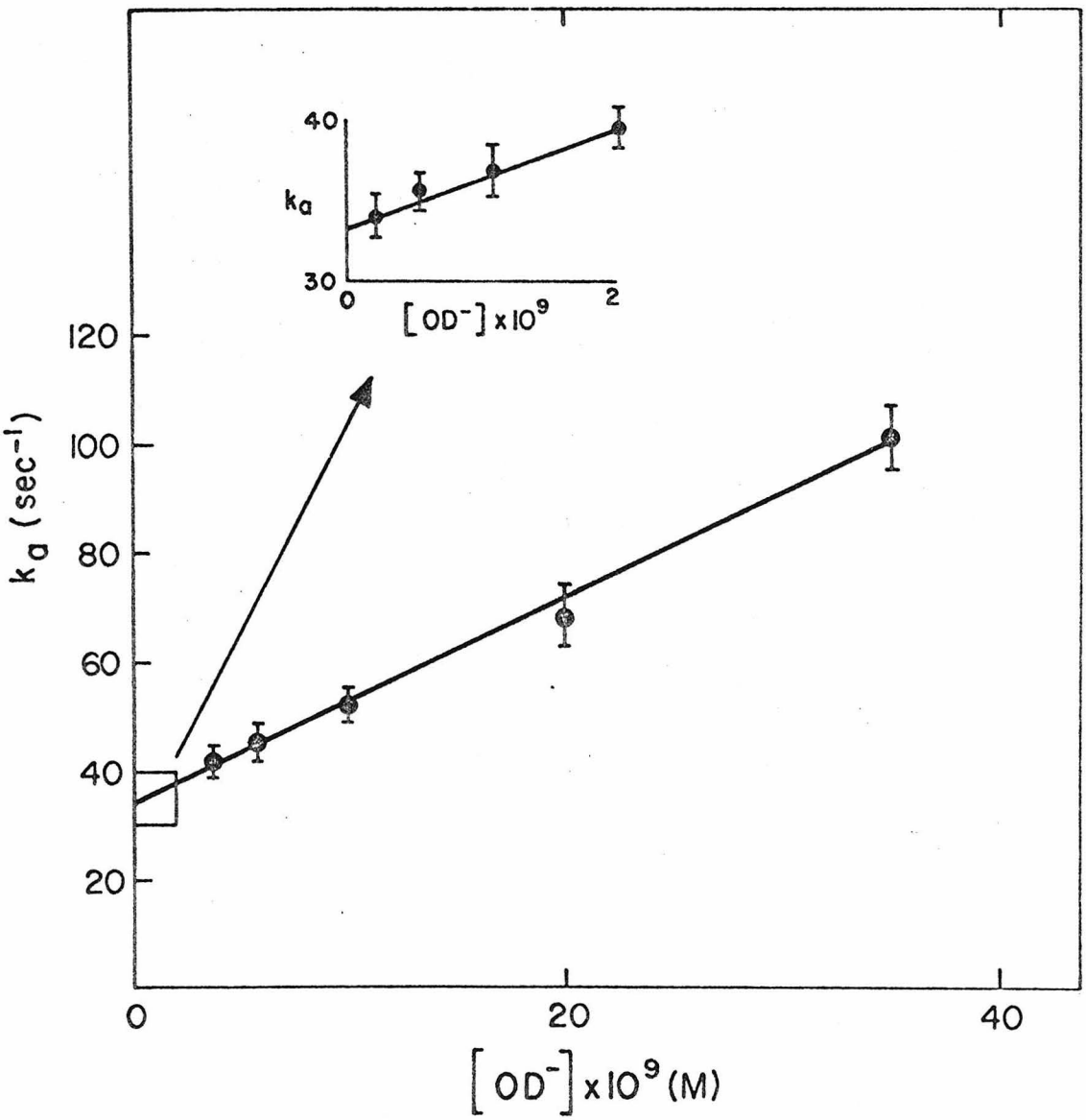
In Fig.25, we have plotted the rate constant extracted from the linewidth data versus the OD^- concentration. (For reasons to be discussed below, we have neglected the linewidth data below pD $\sim$ 3.0.) This plot clearly depicts the involvement of two kinetic pathways and enables us to determine the relative contribution of the solvent and base catalyzed steps. The following rate law at 30°C is deduced from this analysis:

$$-\left(\frac{d[A]}{dt}\right)_f = 1.7 \times 10^9 [OD^-][A] + 0.7 [D_2O][A] \ (M\ sec^{-1}) .$$

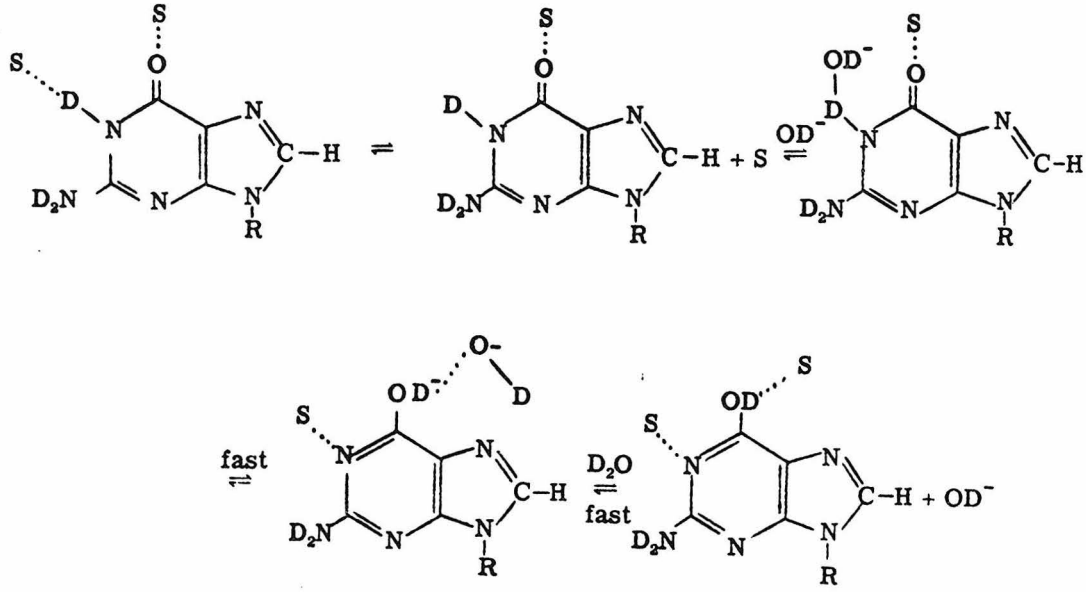
The Arrhenius plot in Fig.24 yielded activation energies of 7 ± 1 kcal/mole and 13 ± 1 kcal/mole. A careful study of the temperature dependence of the H_8 linewidth at several pD's indicates that the lower activation energy component is the pD dependent pathway.

The above rate law suggests the following mechanism for the lactam-lactim tautomeric exchange of the guanine base.

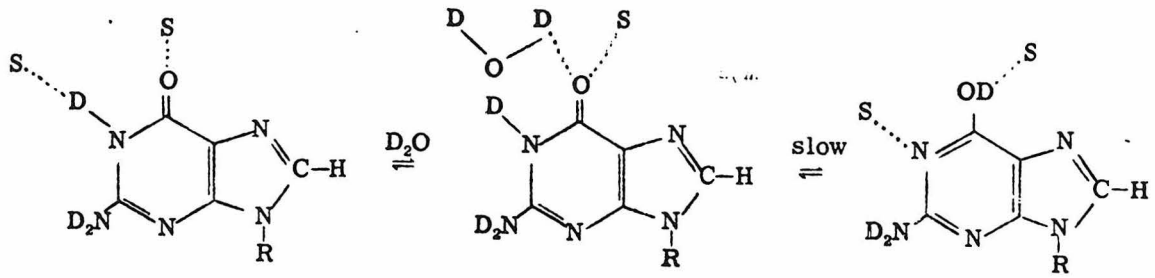
Figure 25. Dependence of the pre-exchange rate constant (k_a) on the base (OD^-) concentration.


40
k_a
30
0
2
$[OD^-] \times 10^9$
120
100
80
60
40
20
k_a (sec^{-1})
0
20
40
$[OD^-] \times 10^9$ (M)

(i) For the base-catalyzed step:



(ii) For the solvent catalyzed step:



Both k_B and k_S' are of the proper order of magnitude for proton abstraction from an ND group by OD^- and D_2O.

As we discussed earlier, the abrupt sharpening of the H_8 linewidth below pD ~ 3.0 is most probably due to the protonation of the guanine base ($pK_a \sim 2.5$). Although the presence of acid can catalyze the exchange between the two tautomeric forms, our linewidth data do not support this. Such an acid-catalyzed mechanism ought to be diffusion controlled, and would lead to a precipitous sharpening of the H_8 resonance at much higher pD's, say pD ~ 6.0. In fact analysis of the low pD data in terms of an acid-catalyzed mechanism yielded a bimolecular rate constant of 2.5×10^4 $M^{-1}sec^{-1}$, a factor of 10^4 lower than that usually accepted for acid-catalyzed processes.(82) We therefore rule out this possibility.

2.3. Effect of Intermolecular Association.

With the exception of 2'-GMP, 2-N,N-dimethylaminomethylene-guanosine and ApG, the H_8 linewidth is concentration dependent over the pD range 3 to 7. The effect of increasing the concentration is one of line narrowing both in the case of 2'-deoxy-guanosine derivatives, which are known not to form gels, and for guanosine, 5'-GMP and 3'-GMP above the melting temperature of the gels. In the case of the latter compounds, gel formation leads to discontinuous broadening of the H_8 resonance, as expected.

In this section we shall be concerned with the equilibrium processes which lead to the observed

sharpening of the H_8 resonance with increasing concentration. We shall assume that the line narrowing arises merely from the effect of intermolecular association on the tautomeric equilibrium, that the association proceeds via formation of the hydrogen-bonded tetramer proposed by Gellert et al.[69], and that under favorable conditions, the tetramer can undergo further aggregation by vertical stacking or other interaction. We shall further assume that tetramer formation is only important when the guanine base is in the normal keto structure, when the guanine base is neutral, and when the phosphate group is monoprotonated in the case of the nucleotides. Finally, it will be assumed in this analysis that the chemical exchange of a guanine base among the various species is rapid compared with the inverse of the shortest nmr transverse relaxation time of the largest molecular aggregate, and that these molecular aggregates do not contribute much broadening to the weighted averaged intrinsic linewidth. As it turned out, this latter approximation is not a serious one, since even at the concentration of 0.01 M 5'-GMP, the highest concentration employed in these studies, the largest molecular aggregates of any significant concentration involves no more than 8 guanine bases and this associated dimer of the tetramer accounts for no more than 13% of the total G molecules in solution.

With these assumptions, the observed H_8

linewidth at a given stoichiometric concentration C_o would be determined by $[G_{lactim}]/C_o$, the fraction of molecules existing in solution in the lactim form irrespective of the degree of ionization of the phosphate group and whether the guanine base is protonated or not. In the absence of extensive intermolecular association, the lactam tautomer should exist predominantly in solution as the monomer. We shall designate this species by G when the guanine base is neutral and its phosphate group is singly ionized. $[G_{lactim}]$ can be related to the equilibrium concontration of G via the following multiple equilibria.

$$G \overset{K_T}{\rightleftharpoons} G^* \tag{20}$$

$$G_- \overset{K_T}{\rightleftharpoons} G_-^* \tag{21}$$

$$G_+ \overset{K_T}{\rightleftharpoons} G_+^* \tag{22}$$

$$G \overset{K_a'}{\rightleftharpoons} G_- + D^+ \tag{23}$$

$$G^* \overset{K_a'}{\rightleftharpoons} G_-^* + D^+ \tag{24}$$

$$G_+ \overset{K_a}{\rightleftharpoons} G + D^+ \tag{25}$$

$$G_+^* \overset{K_a}{\rightleftharpoons} G^* + D^+ \tag{26}$$

In these expressions, the minor lactim tautomer is distinguished from the lactam species by a superscript asterisk, and the subscripts + and - are used to designate the species when the base becomes protonated and when the

phosphate group is doubly ionized respectively. Expressions (20)-(22) denote the tautomeric equilibria between the various lactam and lactim species. Ionization of the secondary hydrogen of the phosphate group is given by expressions (23) and (24); expressions (25) and (26) account for protonation of the guanine base. In these equilibria, we have assumed that the pK_a's of both the guanine base and the phosphate group are not affected by the tautomerism of the base as there is no experimental evidence to suggest the contrary. pH titration of model compounds indicates that the pK_a of the guanine base differs by no more than 0.1 or 0.2 pK units between the lactam and lactim structures. The tautomeric equilibrium is therefore not grossly shifted upon protonation of the base. Ionization of the phosphate group should not affect the lactam-lactim equilibrium.

From the above multiple equilibria, it is readily deduced that

$$[G_{lactim}] = [G_+^*] + [G^*] + [G_-^*]$$

$$= K_T \left(1 + \frac{[D^+]}{K_a} + \frac{K_a'}{[D^+]}\right) [G] \quad . \qquad (27)$$

This equation can be used to ascertain $[G]/C_o$ from the experimentally observed linewidth, if the kinetics of the tautomeric exchange deduced for 2'-GMP can be assumed to be applicable to all the guanine derivatives. Conversely,

the pD and concentration dependence of the H_8 linewidth can be calculated if [G] is known as a function of pD and concentration at a given temperature. Note that in deducing $[G_{lactam}]/C_o$ and hence $[G]/C_o$ from the experimental linewidths, proper considerations must be made to account for changes in the chemical shift difference between the lactam and lactim species with pD and concentration, i.e., changes in the effective nmr time scale. In the discussion to follow, for example, the chemical shift of the G molecules in the molecular aggregates may be somewhat different from that in the unassociated G species, in which case the association process will alter somewhat the effective nmr time scale which determines the chemical exchange contribution to the H_8 linewidth, assuming of course that the interconversion between the lactam and lactim species remains the rate-limiting step in the overall equilibrium. However since the H_8 chemical shift was observed to be independent of the concentration over the complete range of concentration investigated (0.001-0.01 M), the chemical shift difference between G^* and weight averaged G species is changed at most a few Hz over this concentration range, and we may ignore this complication in our analysis.

In our model, intermolecular association decreases the equilibrium concentration of G. Accordingly, $[G_{lactim}]/C_o$ decreases, and the exchange contribution to the H_8 linewidth is reduced. The effect of this molecular

aggregation can be accounted for by including the following processes.

$$4\,G \overset{K_1}{\rightleftharpoons} G_4 \tag{28}$$

$$2\,G_4 \overset{K_2}{\rightleftharpoons} (G_4)_2 \tag{29}$$

$$G_4 + (G_4)_2 \overset{K_3}{\rightleftharpoons} (G_4)_3 \tag{30}$$

$$\vdots \qquad \vdots \qquad \overset{K_n}{\rightleftharpoons} \quad \vdots$$

$$G_4 + (G_4)_{n-1} \rightleftharpoons (G_4)_n \tag{31}$$

Equation (28) gives the equilibrium for tetramer formation. The vertical stacking or association of the tetramer to form higher aggregates is depicted by Eqs. (28) - (31). G_4, $(G_4)_2$, $(G_4)_3$,... $(G_4)_n$ denote the hydrogen-bonded tetramer and its stacked or associated dimer, trimer, n-mer, respectively.

For a given stoichiometric concentration C_o, $[G]$ can be obtained from mass balance considerations. By definition

$$C_o = [G] + [G_+] + [G_-] + [G^*] + [G^*_+] + [G^*_-]$$

$$+ 4\,[G_4] + 8\,[(G_4)_2] + \ldots 4n\,[(G_4)_n] + \ldots, \tag{32}$$

from which it follows that

$$C_o = (1 + K_T)\left(1 + \frac{K_a'}{[D^+]} + \frac{[D^+]}{K_a}\right)[G]$$

$$+ 4K_1[G]^4\{1 + 2[G]^4 K_1 K_2 + \ldots + n[G]^{4(n-1)} K_1^{(n-1)} K_2 K_3 \ldots K_n\}. \tag{33}$$

If we now make the reasonable simplifying assumption that the stacking constants for each successive association step are equal, i.e., $K_2 = K_3 = \ldots . K_n = K$, then equation (33) simplifies to

$$C_o - \eta [G] = \frac{4 K_1 [G]^4}{(1 - K_1 K [G]^4)^2} = f([G]) \quad , \tag{34}$$

$$\text{where } \eta = (1 + K_T)\left(1 + \frac{K_a'}{[D^+]} + \frac{[D^+]}{K_a}\right) \quad , \tag{35}$$

$$\text{and } K_1 K [G]^4 < 1 \quad . \tag{36}$$

It is not difficult to show that the above considerations do reproduce the two-maxima linewidth vs. pD profile experimentally observed for guanosine, 5'-GMP, etc. We shall omit the details here, except to point out that the theory does correctly predict the pD value corresponding to the linewidth minimum. Over the pD range 3 to 6, the kinetics of the tautomerism is essentially independent of the pD, so that the minimum linewidth should occur at a pD for which $[G_{lactim}]/C_o$ is a minimum. Simple differentiation shows that $[D^+]_{min} = \sqrt{(K_a K_a')}$ or $(pD)_{min} = \frac{1}{2}(pK_a' + pK_a) \cong \frac{1}{2}(3+7) = 5$, in agreement with experiment.

We shall now attempt to extract the intermolecular association constants K_1 and K from the linewidth data in the case of 5'-GMP, whose concentration dependence we have studied extensively. To do this, we have resorted to the following graphical method. From Eq. (34), we note that

if $C_o - \eta[G]$ and $f([G]) = 4K_1[G]^4/(1 - K_1K[G]^4)^2$ are plotted versus [G], the curves should intersect at the experimental [G]'s determined from the observed linewidths for one unique value of K_1 and K. Our attempts to fit the experimental data are summarized in Fig. 26 and Fig. 27. In Fig. 26, only the data at low 5'-GMP concentrations have been used. Under these conditions $KK_1[G]^4 << 1$, so that $f(G) \approx 4K_1[G]^4$, and the correct value of K_1 may be determined by comparison of the experimental points with a family of curves of the function $f([G])$ corresponding to different K_1's. In Fig. 26, $f([G])$ is given for K_1 values of 10^7 M^{-3}, 2.5×10^7 M^{-3}, and 10^8 M^{-3}. The curve for $K_1 = 2.5 \times 10^7$ M^{-3} can be seen to best fit the low concentration data of 5'- GMP. After K_1 has been determined, the tetramer stacking or association constant can be obtained by fitting the high concentration data in a similar manner using the full expression for $f([G])$, i.e., without making the $K_1K[G]^4 << 1$ approximation. This graphical determination of K is given in Fig. 27, from which we ascertained that $K = 40 \pm 10$ M^{-1} from the quality of the fit of the high concentration data. The above analysis yields the following distribution of species in solution for 0.01 M 5'-GMP: [G] 27%, $[G^*]$ 5%, $[G_4]$ 53%, $[G_4]_n$ $(n \geq 2)$ 13%.

The analysis contained in this section illustrates the subtle factors which determine the molecular aggregation properties of the guanine nucleosides and nucleotides. Presumably K and K_1 are both small in the case of

Figure 26. Graphical determination of the tetramer formation constant (K_1) for 5'-GMP.

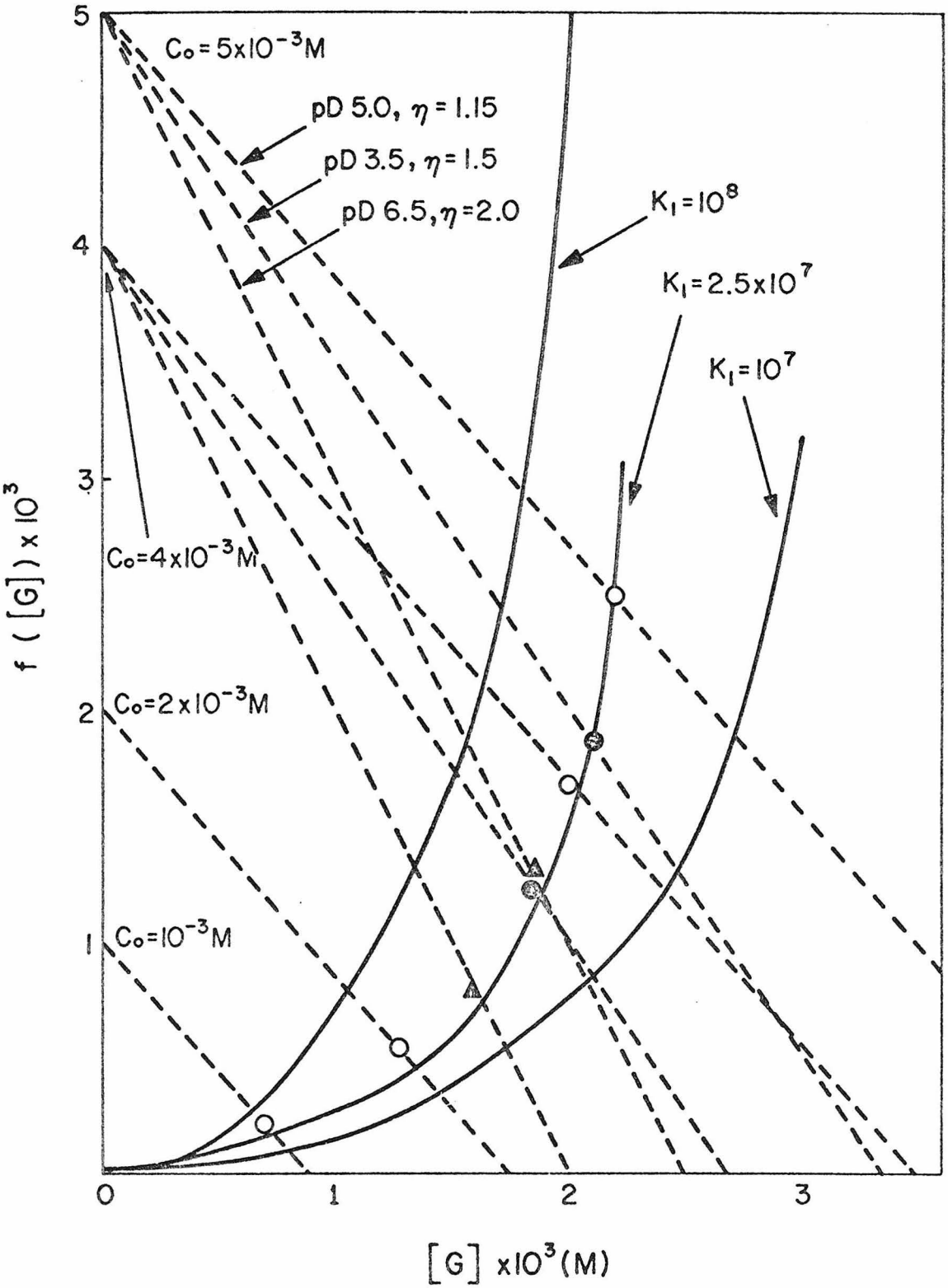
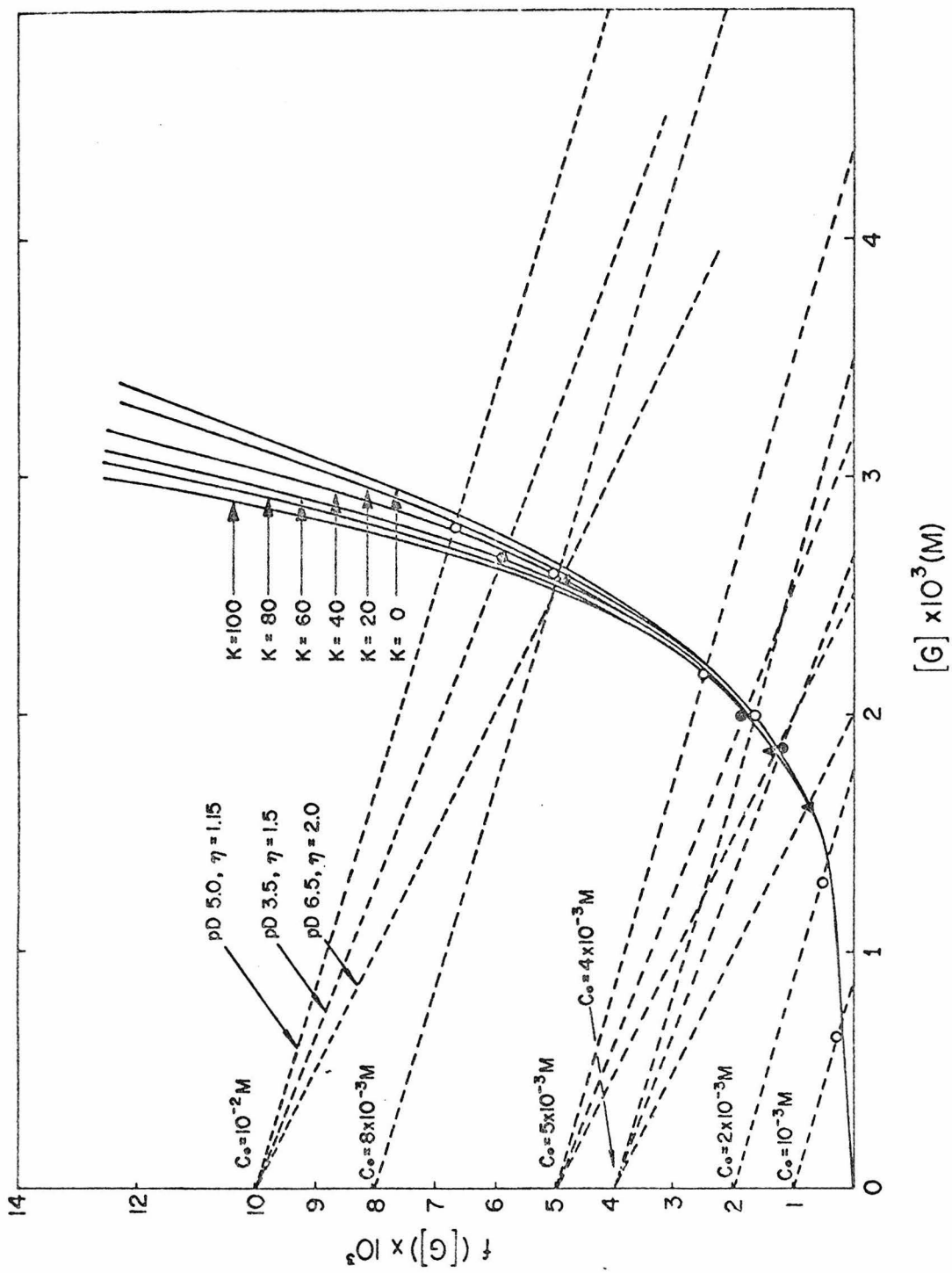

$C_o = 5\times10^{-3}M$
pD 5.0, η = 1.15
pD 3.5, η = 1.5
pD 6.5, η = 2.0
$K_1 = 10^8$
$K_1 = 2.5\times10^7$
$K_1 = 10^7$
$C_o = 4\times10^{-3}M$
$C_o = 2\times10^{-3}M$
$C_o = 10^{-3}M$
f ([G]) x 10^3
[G] x 10^3 (M)
0
1
2
3
4
5

Figure 27. Graphical determination of the tetramer stacking constant (K) for 5'-GMP.


K=100
K= 80
K= 60
K= 40
K= 20
K= 0
pD 5.0, η = 1.15
pD 3.5, η = 1.5
pD 6.5, η = 2.0
C₀=10⁻²M
C₀=8x10⁻³M
C₀=5x10⁻³M
C₀=4x10⁻³M
C₀=2x10⁻³M
C₀=10⁻³M
t ([G]) x 10²
[G] x10³(M)

2'-GMP, 2-N, N-dimethylamino-methylene-guanosine and ApG. In the case of 2'deoxyguanosine, 5'-dGMP and 5'-GTP, the value of K_1 is probably similar to that for guanosine and 5'-GMP, but K is apparently much smaller. Clearly gel formation is possible only when the tetramer stacking constant is large.

3. EDTA as a Catalyst for the Tautomeric Exchange.

3.1. Dramatic Effect of EDTA.

Experimentally, we have observed that the addition of 10^{-6} M EDTA is sufficient to sharpen the broad H_5 and H_8 resonances in cytosine and guanine derivatives respectively. There are two possible explanations for this remarkable phenomenon. The first of this is that the broadening is due to paramagnetic ion contaminants present at a concentration level of 10^{-6}M, and the added EDTA is merely serving as a chelating agent. We have, however, argued that it is unlikely that paramagnetic ion contaminants at this concentration level should lead to the line-broadening observed. Moreover, if this were the case, other efficient chelating agents such as 8-hydroxy-quinoline and ethylene diamine should be effective in collapsing the broadened H_5 and H_8 resonances as well. The fact that these chelating agents do not sharpen the H_5 and H_8 resonances even at concentrations as high as 10^{-3} M would seem to rule out this possibility. This

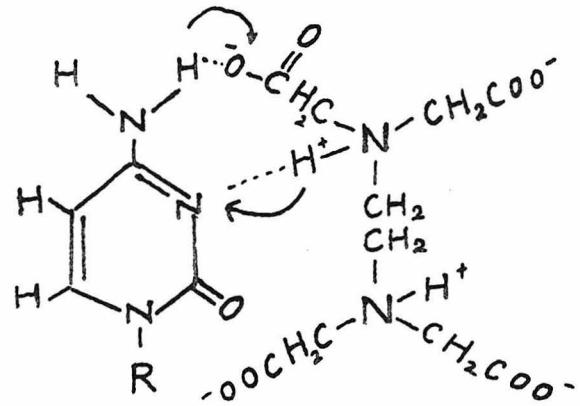
leaves us with the other only alternative, which is that EDTA serves as a catalyst for the tautomeric exchange between the amino and imino tautomers of the cytosine base as well as that between the lactam and lactim tautomers of the guanine base. Since near neutral pD, EDTA is both a proton accepter as well as a proton donor,[83] we suggest that it may act through a concerted mechanism involving : in the case of cytosine simultaneous proton abstraction from the amino group by one of the four carboxylate anions and proton transfer from one of the two protonated EDTA nitrogen sites to the N-3 of the cytosine base (Fig.28a) and in the case of guanine, simultaneous proton abstraction from the N_1-D by one of the four carboxylate anions and proton transfer from one of two protonated EDTA nitrogen sites to the keto group of the guanine base (Fig.28 b). If this mechanism is correct, then the efficiency with which EDTA catalyzes the tautomeric exchange indicates the kinetics is almost diffusion controlled (Fig. 28).

If the effect of EDTA is to catalyze the tautomeric exchange, then we expect $1/\tau = k_B [OD^-] + k_S [D_2O] + k_c [EDTA]$ and at a given pD, $1/\tau$ should vary linearly with [EDTA] . This linear variation is clearly depicted in Fig. 29. From the slope, we deduced a rate constant of $2 \times 10^{+8}$ $M^{-1}sec^{-1}$ for the case of 5'-CMP.

A number of reagents have also been examined to determine if they can serve as catalysts for the tautomeric

Figure 28. Schematic representations for the catalytic mechanism of EDTA on the tautomeric exchange of the cytosine$_{(a)}$ and guanine base$_{(b)}$.

a



b

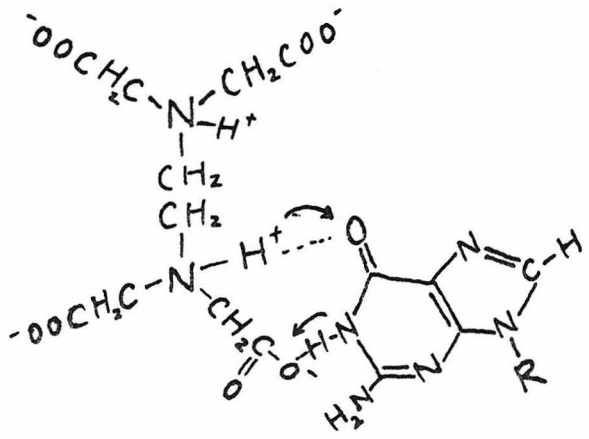
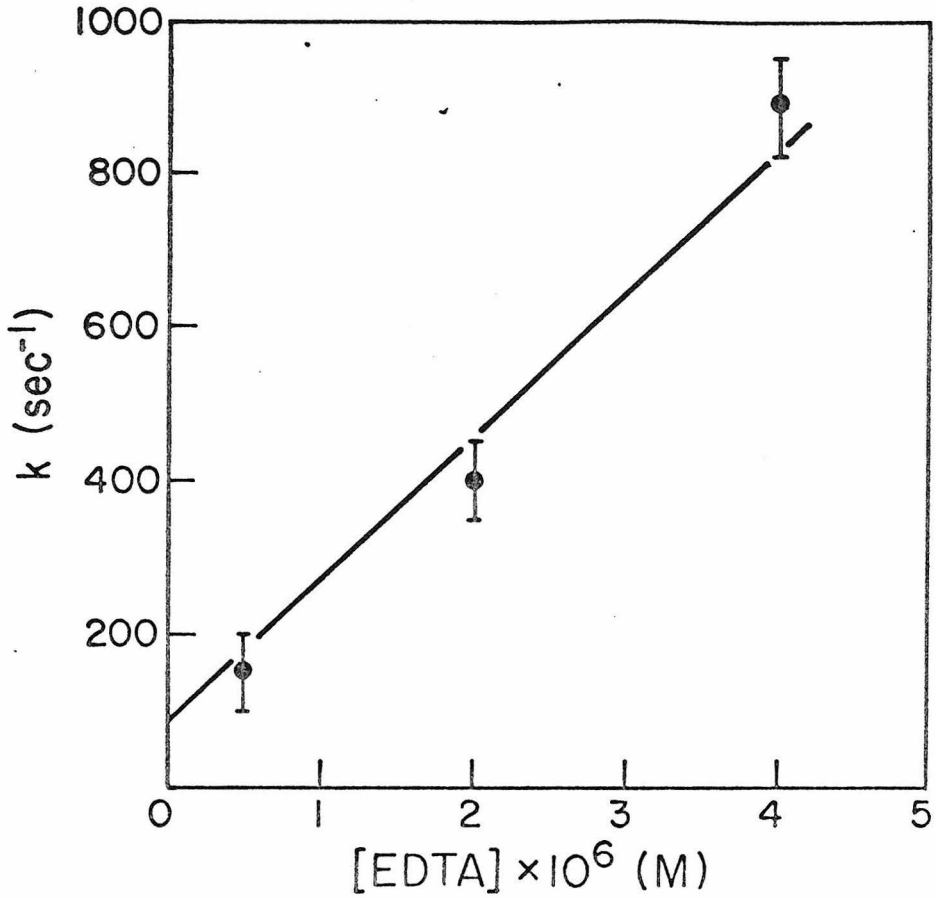
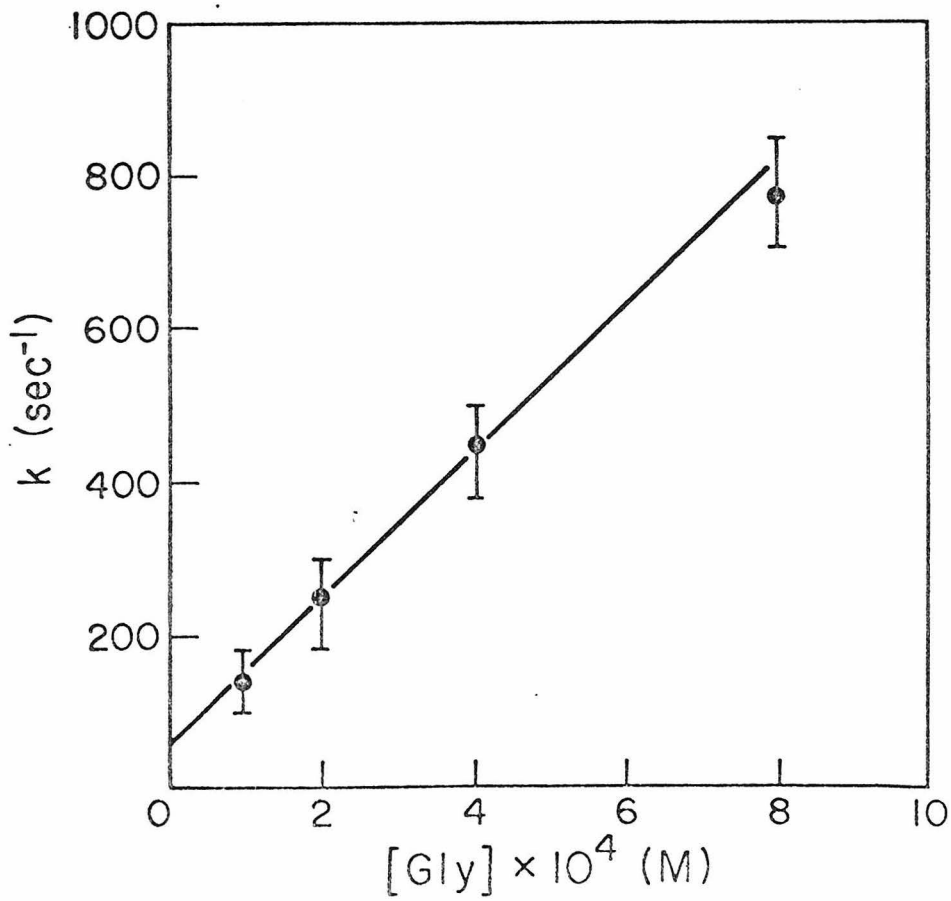


Figure 29. Plots of the tautomeric exchange rate k versus the concentration of EDTA or glycine.


1000
800
600
400
200
k (sec⁻¹)
0
1
2
3
4
5
[EDTA] ×10⁶ (M)


1000
800
600
400
200
k (sec⁻¹)
0
2
4
6
8
10
[Gly] × 10⁴ (M)

exchange of these two bases in solution. A similar study on glycine also showed that the concentration of glycine higher than 10^{-4}M can also affect the linewidth of the H_8 resonance of the guanine base. In view of the low stability constants of glycine-metal complexes (Table IIIA), we conclude that glycine also serves as a catalyst for the tautomeric exchange through a similar mechanism to EDTA. The catalytic rate constant was estimated as $\backsim 10^6$ $M^{-1}sec^{-1}$ --two order of magnitude lower than that from EDTA. This is probably due to the fact that there is only one COO^- in the case of glycine compared to four for EDTA. Also, the proton in the tertiary amine of EDTA should be more labile or acidic compared to that in the primary amine of glycine. For example, the difference in the protolysis-rate constants between the monomethyl ammonium and H_2O and between trimethyl ammonium ions and H_2O are an order of magnitude faster in the trimethyl ammonium,[81] these rate constants being 0.4×10^{-2} and 5.5×10^2 respectively. (in $M^{-1}sec^{-1}$).

Lysine also exhibits the catalytic effect on the tautomeric exchange of the base. However the catalytic rate constant is about one order of magnitude lower than that of glycine, probably due to the steric effect of aliphatic side chain. By contrast NaAc has no measureable effect on the cytosine H_5 linewidth. These experiments demonstrate the requirement of a $-O-\overset{O}{\overset{\|}{C}}-CH_2-\underset{H^+}{N}-$ zwitter ion linkage for the catalyst.

Table of Stability Constants (ⅢA)

	EDTA	ED	Gly	HQ
Co^{2+}	16.2	5.9	5.2	10.6
Cu^{2+}	17.8	10.7	8.6	12.6
Fe^{2+}	13.9	4.4	4.3	8.7
Mn^{2+}	14.0	2.8	3.2	6.8
Ni^{2+}	18.6	7.6	6.0	11.4
Ce^{3+}	15.4			
Dy^{3+}	17.7			9.1
Er^{3+}	17.8			
Eu^{3+}	16.7			La^{3+} 5.9
Fe^{3+}	7.4		10.0	10.0
Lu^{3+}	20.1			
Th^{4+}	23.2			21.0

EDTA: Ethylene diamine tetraacetic acid

ED: Ethylene diamine

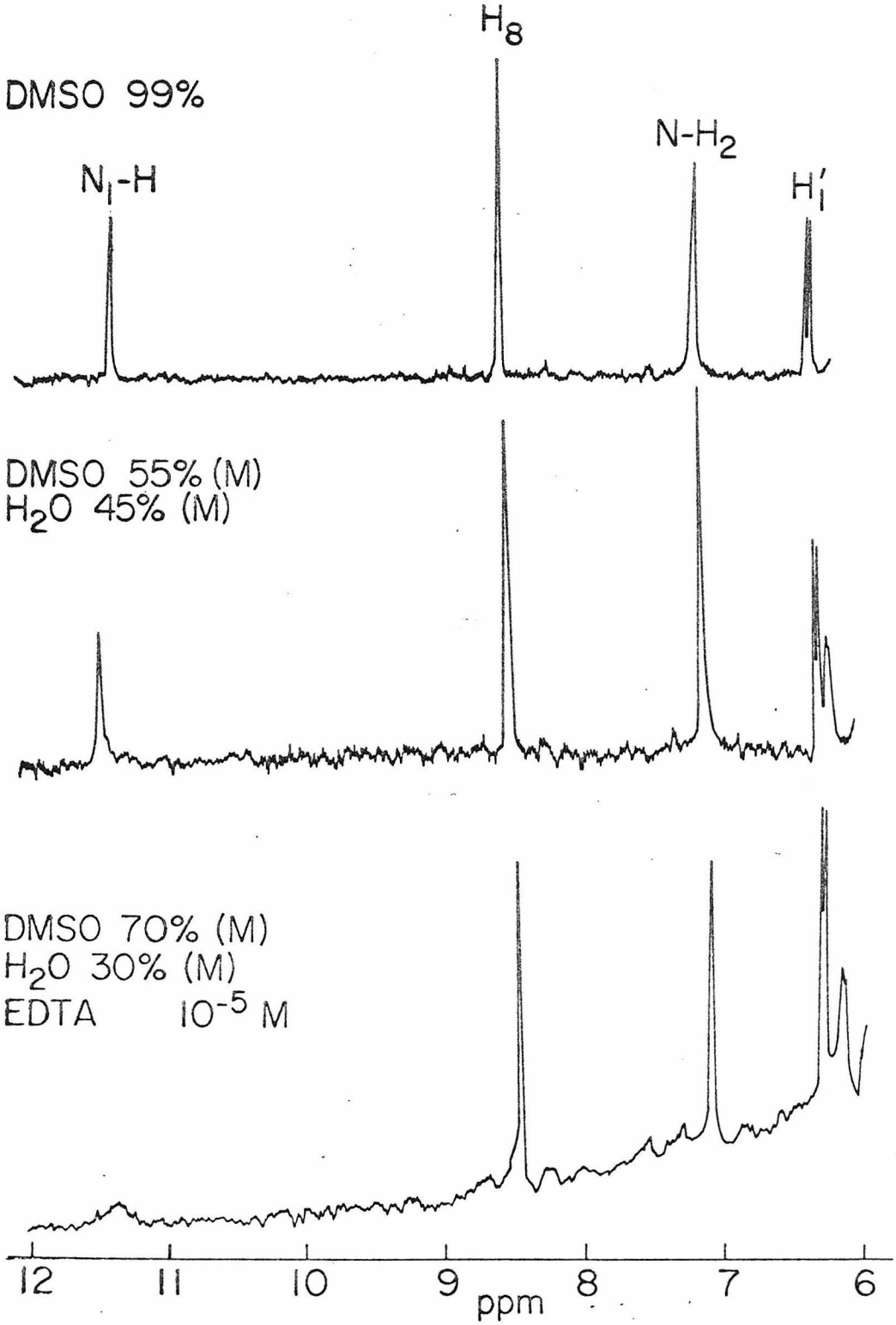
Gly: Glycine

HQ: 8-Hydroxyl quinoline

The above proposed catalytic mechanism for the effect of EDTA is further confirmed by the following experiments in which we attempted to slow down the rate of the abstraction of the NH_2 and NH protons of the bases in DMSO-water mixtures. In dry DMSO, it was observed that the intensities of all the resonances are in direct proportion to the corresponding number of protons in the bases. The cytosine H_5 and guanine H_8 protons appear as sharp resonances whereas the resonances of the NH_2 and NH protons exhibit slight broadening (3 to 5 Hz) probably as a result of the spin-spin coupling of these protons with the quadrupolar ^{14}N nuclei. However, with the addition of 50 mole percent of water into the dry DMSO, the NH_2 and NH proton resonances become further broaded because of the enhanced exchange of these labile protons with those of the water molecules. The resonances of the cytosine H_5 and guanine H_8 protons also exhibit slight broadening in the aforementioned DMSO-water mixture. Furthermore the addition of small traces of EDTA into the mixture was observed to bring about substantial broadening of the proton resonances of the NH_2 group in the cytosine base and of the NH group in the guanine base. The resonances of the cytosine H_5 and guanine H_8 protons were, however, observed to have become sharpened. These results are depicted in Fig. 30 .

These experimental results indicate that in dry DMSO, only one tautomeric form can exist predominantly for each base, i.e., amino form in the case of the cytosine

Figure 30. Effect of EDTA on the various resonances of guanosine in DMSO-water mixtures.


H8
DMSO 99%
N-H2
N1-H
H1'
DMSO 55% (M)
H2O 45% (M)
DMSO 70% (M)
H2O 30% (M)
EDTA 10-5 M
12
11
10
9
8
7
6
ppm

base, and lactam-amino form in that of the guanine base. However, the hydration of the bases in the DMSO-water mixture induced the amino-imino tautomeric exchange of the cytosine base and the lactam-lactim tautomerism of the guanine base. This tautomeric exchange accounts well for the observed broadening of the cytosine H_5 and the guanine H_8 resonances. In the presence of EDTA, the increase in the rate of the abstraction of NH_2 protons in cytosine base, and that of NH proton in guanine base then results in the observed substantial broadening of these resonances. On the other hand, the very presence of EDTA has the effect of catalyzing the tautomeric exchange through the proposed concerted mechanism, thus resulting in the sharpening of the cytosine H_5 and guanine H_8 resonances.

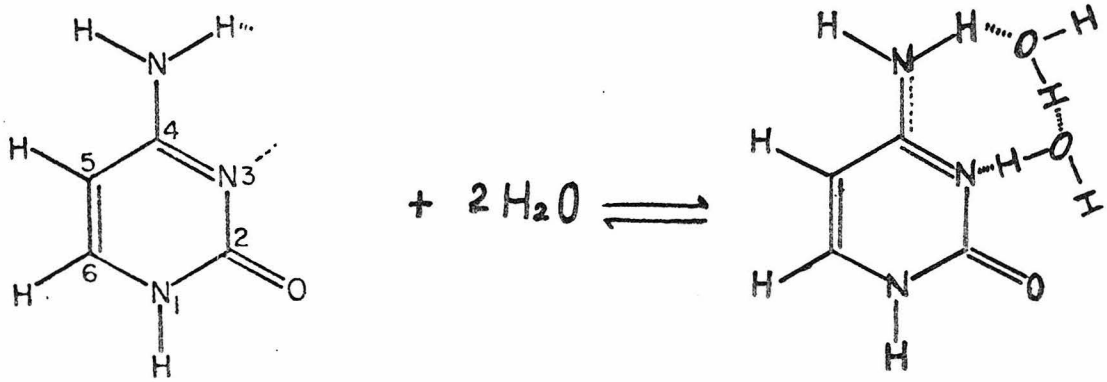
3.2. Effect of Solvent

We have also studied the effect of the solvent on the tautomerism of the nucleic acid bases. In dry DMSO, all the base protons exhibit distinct resonances in the pmr spectra. In the case of guanosine, the H_8 resonance appears as a sharp resonance indicating the absence of paramagnetic ions in the sample. The protons in the NH_2 and NH groups of the base exhibit slightly broad resonances with half widths of 3 and 5 Hz respectively. Judging from the linewidth of the NH proton resonance, we infer that the rate of the proton exchange is quite slow in dry DMSO. No sizeable reduction in the intensity of the H_8

resonance was observed indicating that no minor tautomer exists in appreciable amounts for the guanine base in dry DMSO. However, in the presence of water, the bases become hydrated, and this induces the formation of the minor tautomer through the proton abstraction of the NH group in the base by water molecules. This tautomeric exchange, then gives rise to a slight broadening of the H_8 resonance observed in DMSO-water mixture. From the increase of the linewidth of the NH proton resonance of the guanine base with the addition of water, we may conclude that the rate of the proton abstraction as well as that of the tautomeric exchange is also very slow in DMSO-water mixture.

In the case of cytidine, the amino proton resonances are observed at about 7.5 ppm downfield from TMS. In dry DMSO, the amino protons of this base exhibit a single broad resonance with the half width of 7 Hz, whereas the H_5 and H_6 protons on the same base exhibit sharp resonances. When water is being added to DMSO solution, the original amino proton resonance splits into two, which become progressively farther apart with increasing water concentration. The linewidths of these resonances, however, do not undergo significant changes. The maximum chemical shift difference between these two resonances is about 55 Hz at 220 MHz when the water concentration is 40% (mole). The difference in chemical shifts of these two amino protons in the presence of water indicates that they are

chemically nonequivalent. We believe that these observations are due to the hydration of the cytosine base, which slows down the rotation of the amino group about C-NH_2 bond.



The chemical nonequivalence of the amino protons in the hydrated cytosine base suggests that the partial double bond character of C-NH_2 bond is increased upon the hydration of the base. As a result, the minor imino tautomer may have a stable existence in the presence of water and the water-catalyzed tautomer exchange between the amino and the imino forms of the base may thus proceed via push-pull mechanism. The results of this investigation have hence demonstrated the importance of the base-hydration in the tautomerism of nucleic acid bases.

VI. GUANINE-URACIL BASE-PAIRING

The guanine base in nucleic acid components is known to exist predominantly in the lactam-amino tautomeric structure. In a recent paper, we showed however, that it also exists in appreciable percentages (10-15%) as the abnormal lactim-amino species in aqueous solution at room temperature. Using high resolution proton magnetic resonance (pmr) spectroscopy, we determined both the equilibrium and the kinetics of this lactam-lactim tautomerism. The rate of interconversion between the two tautomers was found to be quite slow, and was dependent on the temperature and pD of the solution. This slow tautomeric exchange was shown to give rise to the guanine H_8 broadening frequently observed in the pmr spectra of guanine derivatives.

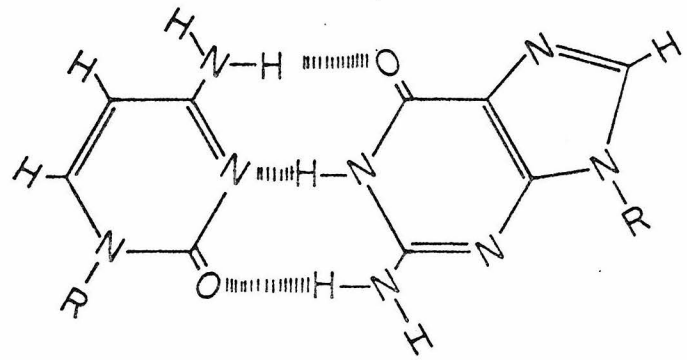
In its normal lactam-amino tautomer, guanine (G) pairs with cytosine (C) to form the well-known Watson-Crick G-C base-pair[3,4] (Fig.31a). This G-C pairing has been observed directly by nmr and ir techniques in a number of organic solvents.[84-87] The lactim tautomer of G, henceforth denoted by G^*, however, has the appropriate electronic structure to pair with an opposing thymine (T) or uracil (U) base (Fig. 31b). Although the lactim tautomer of G is approximately 1 kcal/mole higher in energy than the lactam base, the stability of the abnormal G^* -T or

Figure 31. (a) Normal G-C base-pair.

(b) Abnormal G^{*}-U base-pair.

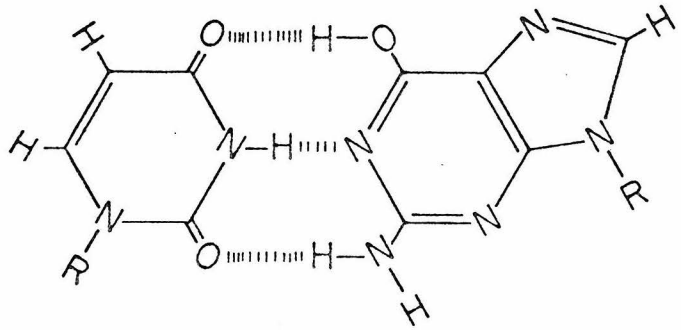
(c) Wobble G-U base-pair.

a



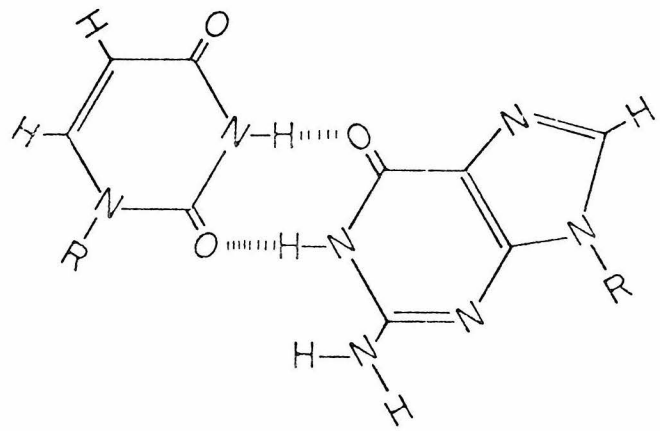
C — G

b



U — G*

c



U — G (Wobble Pairing)

G^*-U base-pair should be comparable to that of the normal G-C base-pair, since the same number of hydrogen-bonds (three) is formed in each case. This consideration suggests that it might be possible to observe G^*-U pairing in solution and to shift the lactam-lactim equilibrium more to the lactim side simply by flooding the solution with U.

We report here a pmr study of this G^*-U base-pairing. Since the lactam-lactim equilibrium was found to be important only when the guanine base is hydrated,[88] it is clear that evidence for G^*-U pairing should be sought in aqueous solution. Unfortunately, such experiments in pure water (H_2O) yield results which are difficult to interpret, both because of the rapid proton exchange between the water protons and the protons which may be involved in the base-pairing, and because of complications arising from the stacking of the nucleic acid bases at the nucleoside or nucleotide concentrations necessary to detect the abnormal base-pairing. Moreover, there has been no evidence for the normal G-C base-pairing in water. It might be that such interactions are quite weak in aqueous solution, except perhaps at low temperatures. We have therefore carried out the G-U base-pairing experiments in DMSO-water mixtures. The water is necessary in order to ensure complete hydration of the guanine base, and shift its lactam-lactim equilibrium far enough toward the lactim side for significant G^*-U pairing to occur. But, by appropriate choice of the water

content, the thermodynamics of the G^*-U base-pairing can be controlled, and the proton exchange rates can be rendered sufficiently slow to permit direct observation of individual resonances for the NH_2 and $N_{(1)}$-H proton of G and the $N_{(3)}$-H proton of U in the pmr spectrum.

The results of our pmr studies of the interaction between guanosine (rG) and 2'-deoxyuridine (dU) are summarized in Tables IV and V. In Table IV, we have summarized the induced chemical shifts which were observed for the $N_{(1)}$-H, NH_2, $H_{(1')}$ and $H_{(8)}$ proton resonances of rG and the $N_{(3)}$-H, $H_{(1')}$, $H_{(5)}$ and $H_{(6)}$ proton resonances of dU for solutions containing 0.1 M rG and 0.2 M dU at various water concentrations in DMSO. Although the shifts observed at this dU/rG concentration ratio are quite small, they are nevertheless large enough to infer that some pairing interaction is occurring between the guanine and uracil bases. The largest downfield shift was observed for that proton resonance which is normally assigned to the $N_{(1)}$-H proton of the guanine base, and significantly, the induced shift observed for this proton increases with the water content from -1.2 to -16.5 Hz over the range of water concentrations studied (3 to 40 mole %). The increased induced shifts observed for the NH_2 and the $N_{(1)}$-H protons of the guanine base with increasing [dU]/[rG] concentration ratio substantiates our contention that G-U pairing is taking place (Table V). In these studies, the water content of the H_2O-DMSO mixture was held fixed at 10 mole%

and the [dU]/[rG] ratio was varied by varying [dU] at fixed [rG].

The downfield shifts observed for the $N_{(1)}$-H and NH_2 protons of rG and the $N_{(3)}$-H proton of dU upon the mixing of these two nucleosides suggest the formation of a hydrogen-bonded complex involving these protons. The absence of a dU induced shift of the amino protons of 1-methyl-guanosine upon the addition of 2'-deoxyuridine to 1-methyl-gaunosine in a 10 mole % H_2O/DMSO mixture indicates no interaction between the bases of these nucleosides when the $N_{(1)}$-H proton of G is replaced by a methyl group or when the $N_{(1)}$-site of G is blocked. Similar experiments with 2-N,N-dimethyl-guanosine and 2'-deoxyuridine also reveal no interaction between these nucleosides, implicating the involvement of the NH_2 group of G in the G-U pairing. Since it is not possible to construct a G-U base-pair involving hydrogen-bonding of both the NH_2 group and the $N_{(1)}$-H proton of G with the guanine base in the lactam-amino structure, we surmise that it is the lactim form of G which is pairing with U. The kinetics of the overall equilibrium is presumably intermediate to fast on the nmr time scale, so that the chemical shift observed for the "$N_{(1)}$-H" proton of G is in actuality a weighted average of the $N_{(1)}$-H proton of the lactam species and the $O_{(6)}$-H proton of the lactim species in both the complexed and the free states. The pairing shifts observed for the G-U interaction may be compared with those observed in a

parallel study of the G-C pairing in pure DMSO (Table IV). Since the lactim tautomer of G is only present to the extent of 10-15% in aqueous solution at room temperature, we expect the base-pairing shifts in the G-U case to be an order of magnitude smaller than those for the G-C case under similar conditions, assuming that the G^*-U and G-C base pairs are of comparable stability. This expectation is clearly borne out by the data summarized in Table I.

The data summarized in Table II can be analyzed semi-quantitatively in terms of the following chemical equilibria:

$$G \overset{K_T}{\rightleftharpoons} G^* \tag{37}$$

$$G^* + U \overset{K}{\rightleftharpoons} G^*\text{-}U \tag{38}$$

Here K_T and K are the quilibrium constants for the lactam-lactim tautomerism and the G^*-U base-pairing respectively. For weak G-U interaction, we expect $[dU_0] \gg [G^*\text{-}U]$ at sufficiently high stoichiometric concentrations of dU, so that the observed base-pairing shift can be adequately approximated by

$$\delta_{obsd} \cong \Delta \frac{K K_T [dU]_0}{(1 + K_T) + K K_T [dU]_0} \tag{39}$$

where Δ is the chemical shift of a proton in the hydrogen-bonded complex relative to that in the uncomplexed state. If we can now assume that Δ for the NH_2 protons of G is

roughly the same in the G^*-U complex as in the G-C complex, i.e., these protons are shifted by about the same amount by G^*-U and G-C pairing, and $K_T \simeq 0.05$ to 0.1, then the data suggest that the equilibrium constant for the pairing between G^* and U is of the order of $1M^{-1}$, which is about the same order of magnitude as that previously reported for the formation of the G-C complex in DMSO. Using this K value and the dU induced shifts observed for the $N_{(1)}$-H ($O_{(6)}$-H) proton of G, we estimate that this proton is shifted $\sim$ 2-3 ppm downfield from its spectral position in the uncomplexed state upon G-U pairing. Since this proton appears 10.5 ppm downfield from TMS in the absence of complex formation, its chemical shift is approximately 13.0 ppm downfield from TMS in the hydrogen-bonded complex, which is what one would expect for the chemical shift of the bridge hydrogen in a relatively strong O-H...O hydrogen-bond.[89]

We have also noted broadening of the N-H proton resonances of G and U upon mixing of the two nucleosides (see Tables IV and V). This linebroadening was found to be more pronounced at higher water contents. At a given water content, the broadening of the $N_{(1)}$-H ($O_{(6)}$-H) resonance of G was also observed to increase with increasing [dU]/[rG] ratio. This water-dependent linebroadening might be due to enhanced proton exchange at the higher water concentrations. However, since we observe little broadening of the N-H

resonances of G and U in comparative experiments involving either rG or dU alone in solution, we exclude this possibility. Thus, the broadening of these resonances in the rG-dU system must be due to the G^{*}-U base-pairing, and we suspect that it arises from chemical exchange, and reflects the dynamics of the overall equilibria denoted by (37) and (38).

Finally, we note that the addition of dU to rG in dry DMSO does not result in significant induced shifts for any of the G or U resonances. This result would seem to suggest that the lactam-lactim equilibrium is not important in DMSO and supports our contention that the lactam-lactim tautomerism is induced by the hydration of the guanine base. The lack of evidence for any form of G-U base-pairing in dry DMSO also enables us to rule out the importance of other base-pairing schemes between G and U when the guanine base is in the lactam-amino form, most notably the "wobble" pairing (Fig.31c) proposed by Crick.[15]

TABLE IV Base-Pairing Shifts[a] in DMSO-H_2O Mixtures

0.1 M rG + 0.2 M dU

H_2O Content Mole %	rG Protons				dU Protons			
	$H_{(1')}$	NH_2	$H_{(8)}$	$N_{(1)}$-H($O_{(6)}$-H)	$H_{(5)}$	$H_{(1')}$	$H_{(6)}$	$N_{(3)}$-H
3	-2.7	+1.2	-2.1	- 1.2 (5 Hz)	-1.6	-1.7	-1.3	+1.4 (6 Hz)
19	-2.6	-1.5	-3.0	- 8.5 (10 Hz)	-2.5	-2.0	-1.0	-3.5 (8 Hz)
31	-2.5	-4.5	-3.5	-10.0 (20 Hz)	-3.4	-2.5	-3.0	-4.0 (16 Hz)
40	-2.5	-3.5	-3.5	-16.5 (30 Hz)	-5.0	-2.0	-3.0	-4.4 (20 Hz)

0.1 M rG + 0.1 M rC

H_2O Content Mole %	rG Protons				rC Protons			
	$H_{(1')}$	NH_2	$H_{(8)}$	$N_{(1)}$-H	$H_{(1')}$	$H_{(5)}$	$H_{(6)}$	NH_2
~ 2	-2	-65.5	-14.0	-144.0	-2.0	-14.0	-12.0	-59.0

[a]Chemical shifts are given in Hz at 220 MHz. Numbers in parentheses denote linewidths (in Hz) of the proton resonance under consideration.

TABLE V G-U Base-Pairing Shifts[a] as a Function of [rG]/[dU] Concentration Ratio in a 10 Mole % H_2O-DMSO Mixture.

[rG], M	[dU], M	rG Protons				dU Protons			
		$H_{(1')}$	NH_2	$H_{(8)}$	$N_{(1)}$-H($O_{(6)}$-H)	$H_{(1')}$	$H_{(5)}$	$H_{(6)}$	$N_{(3)}$-H
0.1	0.2	-2.5	-3.0	-3.5	-5.0 (10 Hz)	+0.5	-1.0	-3.0	-1.5 (10 Hz)
0.1	0.4	-5.0	-6.0	-7.0	-11.5 (25 Hz)	-1.5	-3.0	-4.5	0 (13 Hz)

[a] Chemical shifts are given in Hz at 220 MHz. Numbers in parentheses denote linewidths (in Hz) of the proton resonance under consideration.

VII. CONCLUSION

We have presented experimental evidence to show that the unusual H_5 broadening observed in the pmr spectra of cytosine and its related derivatives as well as the similar H_8 broadening observed in 2'-GMP and related guanine nucleosides and nucleotides at low concentrations is due to chemical exchange between the normal and an abnormal tautomeric structure of these nucleic acid bases. We have studied this line-broadening as a function of temperature, concentration, solution pD, magnetic field as well as for nucleic acid derivatives capable of existing in only the normal or abnormal tautomeric structure. On the basis of a quantitative analysis of the linewidth data, we have ascertained the kinetics of the tautomerism in both systems, have demonstrated that the tautomeric exchange is catalyzed by base (OD^-) as well as the solvent molecules (D_2O), and have concluded that the abnormal tautomer in both the cytosine and guanine systems (imino in the case of cytosine; lactim in the case of guanine) is presented to the extent of 10-15% at room temperature in neutral aqueous solution.

The results of this pmr study indicate that the cytosine base and the guanine base exist in significant percentages in the abnormal tautomeric forms. This result is contrary to what is generally accepted, as all previous investigations of the cytosine and guanine tautomerisms had

provided estimates which are significantly lower than our present indications. From uv and pKa studies of cytosine and its derivatives, Kenner et al.,[23] estimated in 1955 that K_T = [amino] / [imino] = $10^{4.7}$, indicating that the population of the imino tautomer is negligible under ordinary conditions. A similar conclusion was reached by Miles in 1961, who used ir spectroscopy to elucidate the detailed molecular structure of nucleic acid bases in D_2O solution.[39-42] More recently, however, Brown et al.,[35,36] concluded on the basis of uv and ir measurements that $K_T \approx 25$ (in water) for the amino-imino tautomerism in 5,6-dihydrocytosine, a result which, perhaps, is more in line with our present estimates for cytosine. Similar conclusions have been reached in the case of guanine, which has been shown by ur and ir spectroscopy to exist predominantly in the keto-amino or the lactam form in aqueous solution, although, more recently, the work of Wolfenden[27] indicated that $K_T \approx 1$ in aqueous solution for the lactam lactim equilibrium in inosine. In any case, it should be noted that accurate determination of the amino-imino equilibrium in cytosine as well as the lactam-lactim equilibrium in guanine in these earlier studies is limited by the sensitivity and the resolution of the methods used. Although uv spectroscopy is a highly sensitive method for the characterization of electronic structures of molecules, the spectral similarity of nucleic acid bases and their

alkylated analogs and the serious overlap of component bands makes its application to the elucidation of the tautomeric structures of nucleic acid bases inconclusive and unreliable. Similar difficulties are inherent in the ir spectroscopic studies. The determination of tautomeric equilibria by the comparison of pKa's of nucleic acid bases with alkylated analogs is also complicated by the multiplicity of ionizable groups as well as protonation sites. More importantly, we have recently found that the tautomerism of nucleic acid bases is induced upon the hydration of the bases. Hence one is measuring the energy difference between the hydrated species rather than the nonhydrated forms of the bases. In the uv and ir spectroscopic studies, the tautomeric equilibrium constants of the nucleic acid bases were determined simply on the basis of the comparison between the spectra of a given base in aqueous solution and its alkylated base analogs frozen in one or the other tautomeric structure. Since the alkylation on the polar groups of the bases changes significantly the hydrating properties of the bases, such comparative studies are based on unsound rationale and thus have inevitably led to erroneous results. The temperature dependence of the uv spectra of the cytosine base observed by Hélène *et al.*,[32] and Johnson *et al.*,[34] could primarily arise from tautomerism of the nucleic acid bases. Since the cytosine base is strongly hydrated in aqueous solution, changes in uv spectra at high temperatures

cannot be accounted for by considerations of changes in base hydration alone. On the other hand, the tautomerism of the cytosine base should be extremely sensitive to the nature of the solvent, and it may well be that the observed spectral differences for the nucleoside and nucleotide derivatives reflect both differences in the solvation properties of the bases as well as effects of the solvent on the tautomerism.

In the case of cytosine, our results are not exactly incompatible with the theoretical predictions of Pullman and Pullman [48-52], Löwdin[54] and Bodor et al.,[55] all of whom have predicted that the amino form of the cytosine base is more stable than the imino structure (Table IA). The Hückel MO calculations of Pullman and Pullman [48-52] and Löwdin[54] have shown that the resonance energy of the imino form is 2.1 kcal/mole higher than that of the amino tautomer. More recently, Bodor et al.,[55] predicted on the basis of semi-empirical SCF MO calculations that the difference in the heats of atomization between the two tautomers of cytosine is 2.2 kcal/mole higher in the case of the amino structure. In this work, we have estimated that the enthalpy difference between the two tautomers is $\sim$ 1.1 kcal/mole. Considering that the theoretical calculations are made for the molecules in their vacuum states, and hence have neglected solvation effects, these theoretical results may be considered to be in

reasonable agreement with our experimental results. By contrast, the theoretical situation in the case of guanine is in a much less satisfactory state (Table IB). The earlier calculations of Pullman et al.,[48-52] and Hoffmann et al.,[53] have predicted that lactim form of guanine to be more stable than lactam form. While the more recent calculations of Löwdin[54] and Bodor et al.,[55] predict the lactam structure to be more stable, their calculated energy difference between the lactam and lactim is significantly larger than that which is indicated by the present pmr study.

We need not overemphasize the biological significance of our present findings. On the one hand, the high percentages of the abnormal tautomers of guanine and cytosine can provide an important source of spontaneous mutation.[3,5,8] On the other hand, tautomerism of nucleic acid bases can allow for alternate base-pairing schemes which may be important in stabilizing RNA structures[90,91] as well as enzymatic and nonenzymatic codon-anticodon recognition.[21] The odd tautomers of cytosine and guanine, for example, may well provide a molecular rationale for the widespread occurrence of A-C and G-U base pairs thought to be present in t-RNA structures.

REFERENCES

1. G. Mendel, Journal of Heredity, 42, 1 (1951).
2. J.D. Watson, "Molecular Biology of the Gene" (New York, W.A. Benjamin, Inc., 1970).
3. J.D. Watson and F.H.C. Crick, Nature, 171, 737 (1953).
4. M.H.F. Wilkings, A.R. Stokes and H.R. Wilson, Nature, 171, 4 (1953).
5. J.D. Watson and F.H.C. Crick, Nature, 171, 964 (1953).
6. C.R. Woese "The Genetic Code" (New York, Harper and Row, Publishers, 1967).
7. M.Zalokar, Nature, 183, 1330 (1959).
8. J.D. Watson and F.H.C.Crick, Cold Spring Harbor Symp. Quant. Biol., 18, 123 (1953).
9. H.G. Zachan, Angew. Chem. Internat. Edit. 8, 711 (1969).
10. A. Sadgopal, Adv. Gen., 325 (1968).
11. C.R. Woese, Ph.D. Thesis, California Institute of Technology, 1967.
12. S. Sundharadas, J.R. Katze, D. Soll, W. Koningsberg and P. Lengyel, Proc. Nat. Acad. Sci., 61, 693 (1968).
13. S. Nishimara and I.B. Weinstein, Biochemistry, 8, 832 (1969).
14. M.R. Bernfield and M.W. Nibrenberg, Science, 147, 479 (1965).
15. F.H.C. Crick, J. Mol. Biol., 19, 548 (1966).
16. T.D. Sakore and H.M. Sobell, J. Mol. Biol., 43, 77 (1969).

17. G.P. Kreishman, Ph.D. thesis, California Institute of Technology, 1972.

18. R.V. Wolfenden, J. Mol. Biol., 40, 307 (1969).

19. R.V. Wolfenden and S. Hohf, Fed. Proc., 28, 911 (1969).

20. G.C. Medeiros and G.J. Thomas, Jr., Biochim. Biophys. Acta, 238, 1 (1970).

21. G.P. Kreishman and S.I. Chan, J. Mol. Biol., in press (1972).

22. D.L. Barker and R.E. Marsh, Acta Crystallogr., 7, 1581 (1964).

23. G.W. Kenner, C.B. Reese and A.R. Todd, J. Chem. Soc., 855 (1955).

24. A.R. Katritzky and A.J. Waring, J. Chem. Soc., 1540 (1962).

25. A.R. Katritzky and A.J. Waring, J. Chem. Soc., 3046 (1963).

26. D.M. Brown and J, M. Lyall. Austr. J. Chem., 15, 851 (1962).

27. R.V. Wolfenden, J. Mol. Biol., 40, 307 (1969).

28. D. Shugar and J.J. Fox, Biochim. Biophys. Acta, 9, 199, 369 (1962).

29. J. Ueda and J.J. Fox, J. Amer. Chem. Soc., 85, 4024 (1963).

30. L.B. Clark and I Tinoco, Jr., J. Amer. Chem. Soc., 87, 11 (1965).

31. G.C. Medeiros and J. Thomas, Jr., Biochim. Biophys. Acta, 238, 1 (1971).

32. C. Hélène and P. Douzou, C.R. Acad. Sci. Paris, 259, 3385 and 4853 (1964).

33. H. Morita and S. Nagakura, Theoret. Chim. Acta (Berl.) 11, 279 (1968).

34. W.C. Johnson, Jr., P.M. Vipond and J.C. Girod, Biopolymers, 10, 923 (1971).

35. D.M. Brown and M.J.E. Hewlins, J. Chem. Soc., 2050 (1968).

36. D.M. Brown, Pure and Applied Chemistry, 18, 187 (1969).

37. S.J. Angyal and C.L. Angyal, J. Chem. Soc., 1461 (1952).

38. L.N. Short and H.W. Thompson, J. Chem. Soc., 168 (1952).

39. H.T. Miles, Biochim. Biophys. Acta, 22, 247 (1956).

40. H.T. Miles, Biochim. Biophys. Acta, 27, 46 (1958).

41. H.T. Miles, Proc. Nat. Acad. Sci., 47, 793 (1961).

42. H.T. Miles, Biochim. Biophys. Acta, 30, 324 (1958).

43. C.L. Angell, J. Chem. Soc., 504 (1961).

44. C.D. Jardetzky and O. Jardetzky, J. Amer. Chem. Soc., 82, 222 (1960).

45. J.P. Kokko, J.H. Goldstein and L. Mandell, J. Amer. Chem. Soc. 83, 2909 (1961).

46. E.D. Becker, H.T. Miles and R.B. Bradley, J. Amer. Chem. Soc., 87, 5575 (1965).

47. H.T. Miles, R.B. Bradley and E.D. Becker, Science, 142, 1569 (1963).

48. A. Pullman and B. Pullman, Biochim. Biophys. Acta, 64, 403 (1962).

49. A. Pullman and B. Pullman, Biochim. Biophys. Acta, 75, 269 (1963).

50. B. Pullman and A. Pullman, "Quantum Biochemistry" (Interscience Publishers, New York, 1963).

51. A. Veillard and B. Pullman, Compt. Rend., 253, 2277 (1961).

52. A. Pullman and B. Pullman, J. Theor. Biol., 4, 37 (1963).

53. T.A. Hoffman and J. Ladik, "Advances in Chemical Physics" Vol. 7, Chap. 2 (Interscience Publishers, New York, 1964).

54. P.O. Löwdin, "Advances in Quantum Chemistry" Vol. 2, pp. 213-360 (Academic Press, New York, 1965).

55. N. Bodor, M.J.S. Dewar and A.J. Harget, J. Amer. Chem. Soc., 92, 2929 (1970).

56. M.P. Schweizer, A.D. Broom, P.O.P. Ts'o and D.P. Hollis, J. Am. Chem. Soc., 90, 1042 (1968).

57. B.W. Bangerter, Ph.D. Thesis, California Institute of Technology, 1969 .

58. J.H. Nelson, Ph.D. Thesis, California Institute of Technology 1969 .

59. P.O.P. Ts'o , N.S. Kondo, M.P. Schweizer, and D.P. Hollis, Biochemistry, 8, 997,(1969).

60. A. Abragam, "The Principles of Nuclear Magnetism", (Oxford University Press, 1962).

61. J.A. Pople, W.G.Schneider, and H.J. Bernstein, "High Resolution Nuclear Magnetic Resonance", (McGraw-Hill , New York, 1959), pp. 218-225.

62. J.W. Emsley, J. Feeney and L.H. Sutcliffe, "High Resolution Nuclear Magnetic Resonance Spectroscopy", (Pergamon Press, Oxford, 1965).

63. A. Carrington and A.D. McLachlan, "Introduction To Magnetic Resonance," (Harper and Row, Publishers 1967).

64. I. Wempen, R. Duschinsky, L. Kaplan and J.J. Fox, J. Am. Chem. Soc., 83, 4755 (1961).

65. R. Lumry, E.L. Smith and R.R. Glantz, J. Am. Chem. Soc., 73, 4335, (1951).

66. S.S. Danyluk and F.E. Hruska, Biochemistry, 7, 1038 (1968).

67. J.H. Prestegard and S.I. Chan, J. Am. Chem. Soc., 91, 2843, (1969).

68. Gregory Lee, unpublished data.

69. M. Gellert, M.N. Lipsett and D.R. Davies, Proc. Nat. Acad. Sci., 48, 2013 (1962).

70. H.T. Miles and H. Frazier, Biochim. Biophys. Acta, 79, 216 (1964).

71. G.L. Eichhorn, P. Clark and E.D. Becker, Biochemistry, 5, 245, (1966).

72. N. Bloembergen, J. Chem. Phys., 27, 595 (1957).

73. T.J. Swift and R.E. Connick, J. Chem. Phys., 37, 307 (1962).

74. E.D. Becker, H.T. Miles and R.B. Bradley, J. Am. Chem. Soc., 85, 5575 (1965).

75. B.W. Roberts, B. Lambert and J.D. Roberts, J. Am. Chem. Soc., 87, 5439 (1965).

76. P.O.P. Ts'o and S.I. Chan, J. Am. Chem. Soc., 85, 1289 (1963).

77. P.O.P. Ts'o and S.I. Chan, J. Am. Chem. Soc., 86, 4176 (1964).

78. S.I. Chan, M.P. Schweizer, P.O.P. Ts'o and G.K. Helmkamp, J. Am. Chem. Soc., 86, 4182 (1964).

79. O. Jardetzky, Biopolymers Symp. No. 1. 501 (1964).

80. G. Binsch and D.A. Kleier, "The Computation of Complex Exchange-Broadened NMR Spectra Computer Program DNMR," Quantum Chemistry Program Exchange, 1969.

81. E. Grunwald, A. Loewenstein, and S. Meiboom, J. Chem. Phys., 27, 630 (1957); 641 (1957).

82. M. Eigen and L. DeMaeyer, Chapter XVIII, Relaxation Methods, in Techniques of Organic Chemistry, Vol. VIII, A. Weissberger, Ed. Interscience, (John Wiley and Sons New York, 1963), p. 895.

83. R.J. Kula, D.T. Sawyer, S.I. Chan and C.M. Finley, J. Am. Chem. Soc., 85, 2930 (1963).

84. Y. Kyogoku, R.C. Lord and A. Rich, J. Amer. Chem. Soc., 89, 496 (1967).

85. Y. Kyogoku, R.C. Lord and A. Rich., Science, 154, 518 (1966).

86. L. Katz and S. Penman, J. Mol. Biol., 15, 220 (1966).

87. R.A. Newmark, and C.R. Cantor, J. Amer. Chem. Soc., 90, 5010 (1968).

88. G.C.Y. Lee and S.I. Chan, J. Amer. Chem. Soc., in preparation.

89. G.C. Pimentel and A.L. McClellan, "The Hydrogen Bond," (Freeman, New York, 1960).

90. F.H. Martin, O.C. Uhlenbeck and P. Doty, J. Mol. Biol. 57, 201 (1971).

91. O.C. Uhlenbeck, F.H. Martin and P. Doty, J. Mol. Biol., 57, 217 (1971).

APPENDIX I

Consider the general three-site exchange

$$C \underset{\text{fast}}{\rightleftharpoons} A \underset{\text{slow}}{\rightleftharpoons} B$$

Assuming that the exchange between C and A is very fast, while that between A and B is relatively slow and there is no direct exchange between C and B, we may set the Bloch equations modified by the effect of chemical exchange.

$$\frac{dG_A}{dt} + \alpha_A G_A = -i\omega_1 M_o^a + \frac{G_B}{\tau_b} + \frac{G_C}{\tau_c} \tag{1}$$

$$\frac{dG_B}{dt} + \alpha_B G_B = -i\omega_1 M_o^b + \frac{G_A}{\tau_{ab}} \tag{2}$$

$$\frac{dG_C}{dt} + \alpha_C G_C = -i\omega_1 M_o^b + \frac{G_A}{\tau_{ac}} \tag{3}$$

where G_A, G_B, and G_C represent the complex magnetic moments of these three species respectively, $\alpha_A = \frac{1}{T_{2A}} + \frac{1}{\tau_{ab}} + \frac{1}{\tau_{ac}} - i\Delta\omega_A$,

$\alpha_B = \frac{1}{T_{2B}} + \frac{1}{\tau_b} - i\Delta\omega_B$, $\alpha_C = \frac{1}{T_{2C}} + \frac{1}{\tau_c} - i\Delta\omega_C$,

T_{2A}, T_{2B} and T_{2C} denote their respective transverse relaxation times in the absence of exchange.

τ_a, τ_{ab}, τ_{ac}, τ_b and τ_c are their mean life times of exchange respectively,

where $\frac{1}{\tau_a} = \frac{1}{\tau_{ab}} + \frac{1}{\tau_{ac}}$

$$\Delta\omega_A = \omega_A - \omega, \quad \Delta\omega_B = \omega_B - \omega, \quad \text{and } \Delta\omega_C = \omega_C - \omega.$$

ω_A, ω_B, and ω_C represent the resonant frequency of each species respectively. $\omega_1 = \gamma H_1$ where γ is gyromagnetic ratio of proton and H_1 is the reference field. M_o^a, M_o^b, and M_o^c are their equilibrium magnetizations. At steady state we may set

$\frac{dG_A}{dt} = \frac{dG_B}{dt} = \frac{dG_C}{dt} = 0$. From eq. (1), (2), and (3) we have

$$-\alpha_A G_A + \frac{G_B}{\tau_b} + \frac{G_C}{\tau_c} = i\omega_1 M_o^a \tag{4}$$

$$\frac{G_A}{\tau_{ab}} - \alpha_B G_B + 0 = i\omega_1 M_o^b \tag{5}$$

$$\frac{G_A}{\tau_{ac}} + 0 - \alpha_C G_C = i\omega_1 M_o^c \tag{6}$$

Solving these linear equations, we may calculate G_A, G_B, and G_C.

$$G_A = \frac{i\omega_1 \left[M_o^a (\alpha_B \alpha_C) + M_o^b \left(\frac{\alpha_C}{\tau_b}\right) + M_o^c \left(\frac{\alpha_B}{\tau_c}\right) \right]}{-\alpha_A \alpha_B \alpha_C + \frac{\alpha_B}{\tau_{ac}\tau_c} + \frac{\alpha_C}{\tau_{ab}\tau_b}} = u_A + i v_A \tag{7}$$

$$G_B = \frac{i\omega_1 \left[M_o^{a} \left(\frac{\alpha_C}{\tau_{ab}} \right) + M_o^{b} \left(\alpha_A \alpha_C - \frac{1}{\tau_{ac}\tau_c} \right) + M_o^{c} \left(\frac{1}{\tau_c \tau_{ab}} \right) \right]}{- \alpha_A \alpha_B \alpha_C + \frac{\alpha_B}{\tau_{ac}\tau_c} + \frac{\alpha_C}{\tau_{ab}\tau_b}} = u_B + iv_B \quad (8)$$

$$G_C = \frac{i\omega_1 \left[M_o^{a} \left(\frac{\alpha_B}{\tau_{ac}} \right) + M_o^{b} \left(\frac{1}{\tau_{ac}\tau_b} \right) + M_o^{c} \left(\alpha_A \alpha_B - \frac{1}{\tau_{ab}\tau_b} \right) \right]}{- \alpha_A \alpha_B \alpha_C + \frac{\alpha_B}{\tau_{ac}\tau_c} + \frac{\alpha_C}{\tau_{ab}\tau_b}} = u_C + iv_C \quad (9)$$

where $M_o^{a} = P_A M_o$; $M_o^{b} = P_B M_o$; $M_o^{c} = P_C M_o$; and P_A, P_B, and P_C are their relative populations.

Since G_A, G_B, and G_C are complex quantity, we can express them respectively in terms of real part and imaginary part; the former is referred to as the in phase or dispersion moment (u_A, u_B, u_C), and the latter is the out-of phase or absorption moment (v_A, v_B, v_C).

If we assume that all the exchanges involved are in the fast exchange limit, i.e.,

$\tau_a^{-1} \gg \Delta\omega_A \gg T_{2A}^{-1}$; $\tau_b^{-1} \gg \Delta\omega_B \gg T_{2B}^{-1}$; $\tau_c^{-1} \gg \Delta\omega_C \gg T_{2C}^{-1}$

and also $\tau_c^{-1} \gg \tau_b^{-1}$.

We may simplify Eqs. (7), (8), and (9) into the following forms after the complicated calculation and

proper approximations.

$$G_A = \frac{\omega_1 M_o^a \{ [(P_A + P_C)\, \tau_b\, \Delta\omega_B + (P_A + P_B)\, \tau_C\, \Delta\omega_C\ ] + i \}}{\{ [\pi \Delta \nu_o + \frac{P_B^2}{P_A} (1 - P_B)^2\, (\omega_{AC} - \omega_B)^2\, \tau_{ab}\ ] + i\,(\omega_o - \omega) \}}$$

$$= \frac{\omega_1 M_o^a \{ [(P_A + P_C)\, \tau_b\, \Delta\omega_B + (P_A + P_B)\, \tau_C\, \Delta\omega_C\ ] + i \}}{\{ (\frac{1}{T_2})_{obs} + i\,(\omega_o - \omega) \}} \qquad (10)$$

where $\Delta\nu_o = \frac{P_A}{T_{2A}} + \frac{P_B}{T_{2B}} + \frac{P_C}{T_{2C}}$ is the observed linewidth in the absence of exchange.

$\omega_o = P_A\omega_A + P_B\omega_B + P_C\omega_C$ is the observed resonant frequency.

$\omega_{AC} = \frac{P_A\omega_A + P_C\omega_C}{P_A + P_C}$ is the average resonant frequency between species A and C.

$(T_2)_{obs}$ is the observed tansverse relaxation time where

$$(\frac{1}{T_2})_{obs} = \pi\nu_o + \frac{P_B^2}{P_A} (1 - P_B)^2\, (\omega_{AC} - \omega_B)^2\, \tau_{ab} \qquad (11)$$

represents the observed linewidth.

The component of the complex moment which is important to our NMR study is its imaginary one (ν_A, ν_B, ν_C)

which can be obtained from the rationalization of the complex moment.

$$\nu_A = \frac{\omega_1 M_o^a \left[\left(\frac{1}{T_2}\right)_{obs} - (\omega_o - \omega)\ (P_A + P_C)\ \tau_b \Delta\omega_B \right]}{(\omega_o - \omega)^2 + \left[\left(\frac{1}{T_2}\right)_{obs} \right]^2} \qquad (12)$$

Now absorption moment ν_A is simplified as a standard for of Lorentz line shape. By similar mathematical treatment we can get G_B and G_C from which we can calculate ν_B and ν_C, i.e.;

$$B = \frac{\omega_1 M_o^b \left[\left(\frac{1}{T_2}\right)_{obs} - (\omega_o - \omega)\ \tau_b\ (P_A \Delta\omega_A + P_C \Delta\omega_C) \right]}{(\omega_o - \omega)^2 + \left[\left(\frac{1}{T_2}\right)_{obs} \right]^2} \qquad (13)$$

$$C = \frac{\omega_1 M_o^c \left[\left(\frac{1}{T_2}\right)_{obs} - (\omega_o - \omega)\ \tau_b\ \Delta\omega_B (P_A + P_C) \right]}{(\omega_o - \omega)^2 + \left[\left(\frac{1}{T_2}\right)_{obs} \right]^2} \qquad (14)$$

If the exchanges are fast, only one resonance will appear in ω_o which would be properly represented by $\nu_{obs} = \nu_A + \nu_B + \nu_C$ in Eq. (15).

$$\nu_{obs} = \frac{\omega_1 M_o \left\{ \left(\frac{1}{T_2}\right)_{obs} - (\omega_o - \omega)\ \tau_b\ (P_A + P_C) \left[(P_A + P_C) \Delta\omega_B + P_B \Delta\omega_{AC} \right] \right\}}{(\omega_o - \omega)^2 + \left[\left(\frac{1}{T_2}\right)_{obs} \right]^2} \qquad (15)$$

When the exchange rate between A and B is slow, i.e.; $\tau_b^{-1} << \Delta\omega_B$, but that between A and C still remains fast, then the resonance of B will be separated from that of A and C. By approximation, we may calculate the complex moment and relaxation time of each resonance.

$$G_A = \frac{\omega_1 M_o^a \left[(P_A + P_C)\,\Delta\omega_B + i\,\tau_b^{-1} \right]}{\left[(\omega_{AC} - \omega) + i(\frac{1}{T_2})_{AC}\right]\left[(\omega_B - \omega) + i(\frac{1}{T_2})_B\right]} \tag{16}$$

$$G_B = \frac{\omega_1 M_o^b \left[(P_A + P_B)\,\Delta\omega_{AC} + i\,\tau_b^{-1} \right]}{\left[(\omega_{AC} - \omega) + i(\frac{1}{T_2})_{AC}\right]\left[(\omega_B - \omega) + i(\frac{1}{T_2})_B\right]} \tag{17}$$

$$G_C = \frac{\omega_1 M_o^c \left[(P_A + P_C)\,\Delta\omega_B + i\,\tau_b^{-1} \right]}{\left[(\omega_{AC} - \omega) + i(\frac{1}{T_2})_{AC}\right]\left[(\omega_B - \omega) + i(\frac{1}{T_2})_B\right]} \tag{18}$$

where

$$(\frac{1}{T_2})_{AC} = \pi\,\Delta\nu_{oAC} + \tau_{ab}^{-1} \tag{19}$$

$$(\frac{1}{T_2})_B = \pi\,\Delta\nu_{oB} + \tau_b^{-1} \tag{20}$$

where $\pi\Delta\nu_{oAC} = \frac{1}{P_A + P_C}\left(\frac{P_A}{T_{2A}} + \frac{P_C}{T_{2C}}\right)$ and $\pi\Delta\nu_{oB} = \frac{1}{T_{2B}}$ represent the natural linewidth. $(T_2)_{AC}$ is the relaxation

time for the resonance AC and $(T_2)_B$ is that for the resonance B. Thus $(\frac{1}{T_2})_{AC}$ and $(\frac{1}{T_2})_B$ represent their respective linewidths. By suitable treatment, we may separate the total moment $G = G_A + G_B + G_C$ into two terms which represent the moment of the resonances of AC and B respectively. Its imaginary part can also be separated into two terms, i.e.,

$$\nu_{obs} = \nu_A + \nu_B + \nu_C = \frac{\overline{X}}{(\omega_{AC} - \omega)^2 + [(\frac{1}{T_2})_{AC}]^2} + \frac{\overline{Y}}{(\omega_B - \omega)^2 + [(\frac{1}{T_2})_B]^2}$$

where $\overline{X}$ and $\overline{Y}$ are complicated functions for these parameters.

From the absorption moment ν_{obs}, we might understand that there are two separate resonances which occur at frequency ω_{AC} and ω_B when the exchange rate of A and B is slow compared to their chemical shift difference.

Density Matrix Method

Because of the spin-spin coupling with the neighboring H_6 proton, the H_5 resonance was observed as a doublet. Since it belongs to an A X type spectrum, the simple treatment described above is adequate to predict its relaxation time T_2 or its linewidth. However, a more elegant and complete treatment by density matrix method is necessary in the computer simulation of its lineshape.

For the general three sites exchange in the presence of spin coupling with neighboring protons, the similar expression in terms of density matrix can be derived as the following:

$$i\dot{\rho}^{a}_{13} = [H_A, \rho^A]_{13} - i\tau_a^{-1}\rho^A_{13} + i\tau_{ac}^{-1}\rho^C_{13} + i\tau_{ab}^{-1}\rho^B_{13} \quad \text{---------- (1)}$$

$$i\dot{\rho}^{b}_{13} = [H_B, \rho^B]_{13} - i\tau_b^{-1}(\rho^B_{13} - \rho^A_{13}) \quad \text{---------- (2)}$$

$$i\dot{\rho}^{c}_{13} = [H_C, \rho^C]_{13} - i\tau_c^{-1}(\rho^C_{13} - \rho^A_{13}) \quad \text{---------- (3)}$$

$$i\dot{\rho}^{a}_{24} = [H_A, \rho^A]_{24} - i\tau_a^{-1}\rho^A_{24} + i\tau_{ac}^{-1}\rho^C_{24} + i\tau_{ab}^{-1}\rho^B_{24} \quad \text{---------- (4)}$$

$$i\dot{\rho}^{b}_{24} = [H_B, \rho^B]_{24} - i\tau_b^{-1}(\rho^B_{24} - \rho^A_{24}) \quad \text{---------- (5)}$$

$$i\dot{\rho}^{c}_{24} = [H_C, \rho^C]_{24} - i\tau_c^{-1}(\rho^C_{24} - \rho^A_{24}) \quad \text{---------- (6)}$$

where $\dot{\rho}^{s}_{ij}$ represent the components of the density matrix per molecule of H_5 proton of respective species H's are the Hamiltonians of H_5 and H_6 protons of these three species.

By Fourier-analyzing the frequency component of these density matrix equations, we may get the following simple set of equations which are identical to those described in simple case

$$\begin{pmatrix} \alpha_A^{13} & \tau_{ab}^{-1} & \tau_{ac}^{-1} \\ \tau_b^{-1} & \alpha_B^{13} & 0 \\ \tau_c^{-1} & 0 & \alpha_C^{13} \end{pmatrix} \cdot \begin{pmatrix} \rho_{13}^{a(-\omega)} \\ \rho_{13}^{b(-\omega)} \\ \rho_{13}^{c(-\omega)} \end{pmatrix} = -i \frac{\tilde{D}}{2} M_0 \begin{pmatrix} I \\ I \\ I \end{pmatrix}$$

$$\begin{pmatrix} \alpha_A^{24} & \tau_{ab}^{-1} & \tau_{ac}^{-1} \\ \tau_b^{-1} & \alpha_B^{24} & 0 \\ \tau_c^{-1} & 0 & \alpha_C^{24} \end{pmatrix} \cdot \begin{pmatrix} \rho_{24}^{a(-\omega)} \\ \rho_{24}^{b(-\omega)} \\ \rho_{24}^{c(-\omega)} \end{pmatrix} = -i \frac{\tilde{D}}{2} M_0 \begin{pmatrix} I \\ I \\ I \end{pmatrix}$$

where $\alpha_A^{13} = (T_{2A}^{-1} + \tau_{ab}^{-1} + \tau_{ac}^{-1}) - i(\omega_A + \frac{J_A}{2} - \omega)$

$$\begin{cases} \alpha_B^{13} = (T_{2B}^{-1} + \tau_b^{-1}) - i(\omega_B + \frac{J_B}{2} - \omega) \\ \alpha_C^{13} = (T_{2C}^{-1} + \tau_c^{-1}) - i(\omega_C + \frac{J_C}{2} - \omega) \\ \alpha_A^{24} = (T_{2A}^{-1} + \tau_{ab}^{-1} + \tau_{ac}^{-1}) - i(\omega_A - \frac{J_A}{2} - \omega) \\ \alpha_B^{24} = (T_{2B}^{-1} + \tau_b^{-1}) - i(\omega_B - \frac{J_B}{2} - \omega) \\ \alpha_C^{24} = (T_{2C}^{-1} + \tau_c^{-1}) - i(\omega_C - \frac{J_C}{2} - \omega) \end{cases}$$

Assuming $J_A \approx J_B \approx J_C$ = coupling constants between H_5 and H_6 protons

The lineshape function of H_5 resonance can be properly represented by the following.

$$L(-\omega) = P_A (\rho_{13}^{a(-\omega)} + \rho_{24}^{a(-\omega)}) + P_B (\rho_{13}^{b(-\omega)} + \rho_{24}^{b(-\omega)})$$

$$+ P_C (\rho_{13}^{c(-\omega)} + \rho_{24}^{c(-\omega)})$$

We can simulate the lineshape of H_5 resonance by solving the above two sets of linear equations using known parameters at various frequencies.

PART II

PROTON MAGNETIC RESONANCE STUDIES OF HISTONES AND A DNA-HISTONE COMPLEX

I. SOLUTION PROPERTIES OF NUCLEOPROTEINS- HISTONE I AND HISTONE IIb1

1. INTRODUCTION

Nucleoproteins, histone I and histone IIb1 are characterized by their high content of basic amino acid residues, such as lysine and arginine and are known to complex with native DNA to repress the template or gene activity by inhibiting the RNA transcription[1-3]. These proteins are essentially random coil polypeptides in neutral salt free aqueous solution. However at high protein concentrations, in the presence of salt or at extreme pD's it has been shown that these proteins partially change their solution conformations and aggregate by intramolecular and/or intermolecular associations. Qualitative studies about the solution properties of these proteins have been carried out by various workers using optical rotatory dispersion (ORD)[4], circular dichroism (CD)[4], sedimentation equilibrium[5] as well as proton magnetic resonance spectroscopy (pmr)[6-9]. However quantitative studies of the protein aggregations are difficult and far from satisfactory. This is primarily due to the complexity of the system and the lack of a unique and reliable method to monitor the conformation of these molecules.

Experimentally we observed that the spectral area of the pmr spectra corresponding to various resonances of these proteins is

sensitive to either the protein concentrations, the presence of salt or solution pD's. On the basis of the known nmr theory, it can be shown that this phenomenon is closely related to the aggregation of these proteins. From the detailed studies of the spectral area changes under various experimental conditions, we have been able to ascertain the mechanisms as well as the thermodynamics of the aggregations of these proteins.

2. EXPERIMENTAL

Histone I and histone IIb1 kindly provided by Dr. Bonner at the California Institute of Technology were isolated from calf thymus tissue by the selective solvent extraction method of Johns[10]. The purity of these two proteins was checked on acrylamide gel electrophoresis columns. Histone I and histone IIb1 solutions were prepared in 99.7% D_2O supplied by Columbia Organic Chemicals Co., Inc. High field region spectra of histone I and histone IIb1 were obtained with a Varian HR-220 pmr spectrometer equipped with a frequency sweep programmer as well as a Varian C-1024 time averaging computer. A capillary containing 2,2-diphenyl-1-picrylhydrazine (DPPH), tetramethylsilane (TMS) and $CDCl_3$ was used as a spectral area as well as a chemical shift reference. The reported chemical shifts are referred to the internal 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) which has been calibrated with external TMS capillary. Because of the uncertainty in the spectral base line,

the accuracy of this spectral area measurement is about ± 10% depending on the resolution of each spectral region. The total proton concentrations corresponding to the observed spectral area were obtained by comparing with DPPH broadened TMS capillary which has been calibrated with known concentration of internal $(CH_3)_4NCl$. All the spectra were recorded at the ambient probe temperature of 17°C except for the temperature study where the variable temperature controller was used and ethylene glycol was used as a temperature reference. The salts, NaCl, $NaClO_4$, and $MgCl_2$ were of reagent grade. The pD of each solution was measured with a Leeds and Northrup 7401 pH meter equipped with miniature electrodes and was taken to be the observed pH meter reading plus 0.4[11]. Where adjustments were necessary, small amounts of 1 M DCl or 1 M NaOD were added.

3. RESULTS

(1). Experimental and Calculated PMR Spectra

The pmr spectra of histone I and histone IIb1 have been previously reported. A comparison between the experimental and calculated spectra based on the model of a random coil polypeptide[12] is shown in Figs. 1 and 2. The intensities and chemical shifts of the side chain protons of the amino acid residues of these two proteins are represented by the height and the position of the vertical line at the bottom of the calculated spectra. As shown in Figs. 1 and 2, both the chemical shifts

and the spectral width of the experimental spectra agree qualitatively well with the calculated ones except that in some spectral regions, the observed spectra are somewhat broader than the calculated ones. This comparison indicates that histone I and histone IIb1 are essentially random coils in neutral salt free aqueous solution. 100 MHz Fourier transform spectra of histone I and histone IIb1 are given in Fig. 3.

For the convenience of our study, we have divided the high field spectra into several spectral regions. Region I corresponds to the resonances of the methyl protons of some aliphatic amino acid residues, such as Ile, Leu, Val, Region II to those of threonine and leucine residues, Region III and Region IV to those of the side chain protons of lysine, arginine as well as analine residues. Region V corresponds to βCH_2 group of glutamic acid and proline residues. A sharp peak in this region was identified as αCH_3 group of acetate contaminants of the sample. Region VII corresponds to the part of the side chain protons of proline, valine, glutamic as well as aspartic acid residues. Regions VII and VIII are mainly those of the terminal CH_2 group of the side chains of lysine as well as arginine residues. A brief classification of the spectral regions and their corresponding amino acid residues as well as the expected percentages of the proton concentration observed is given in Table IA and IB.

TABLE IA

Spectral regions and their corresponding proton resonances as well as percentages of the total proton concentration of histone I

HIS I

Spectral Region	Resonances Observed	% of Total Proton Concentration
I	Ile - CH_3 Leu - CH_3 Val - CH_3	15%
II	Thr - CH_3 Ala - CH_3 Leu - CH_2	5%
III + IV	Ala - CH_3 Lys - β CH_2, γ CH_2, δ CH_2 Leu - βCH_2, γ CH_2 Arg - βCH_2, γ CH_2	50%
V	Glu - β CH_2 Pro - βCH_2, γ CH_2	9%
VI	Val - β CH Glu - γ CH_2 Asp - β CH_2	5%
VII + VIII	Lys - ϵ CH_2 Arg - δCH_2 Pro - δ CH_2	16%

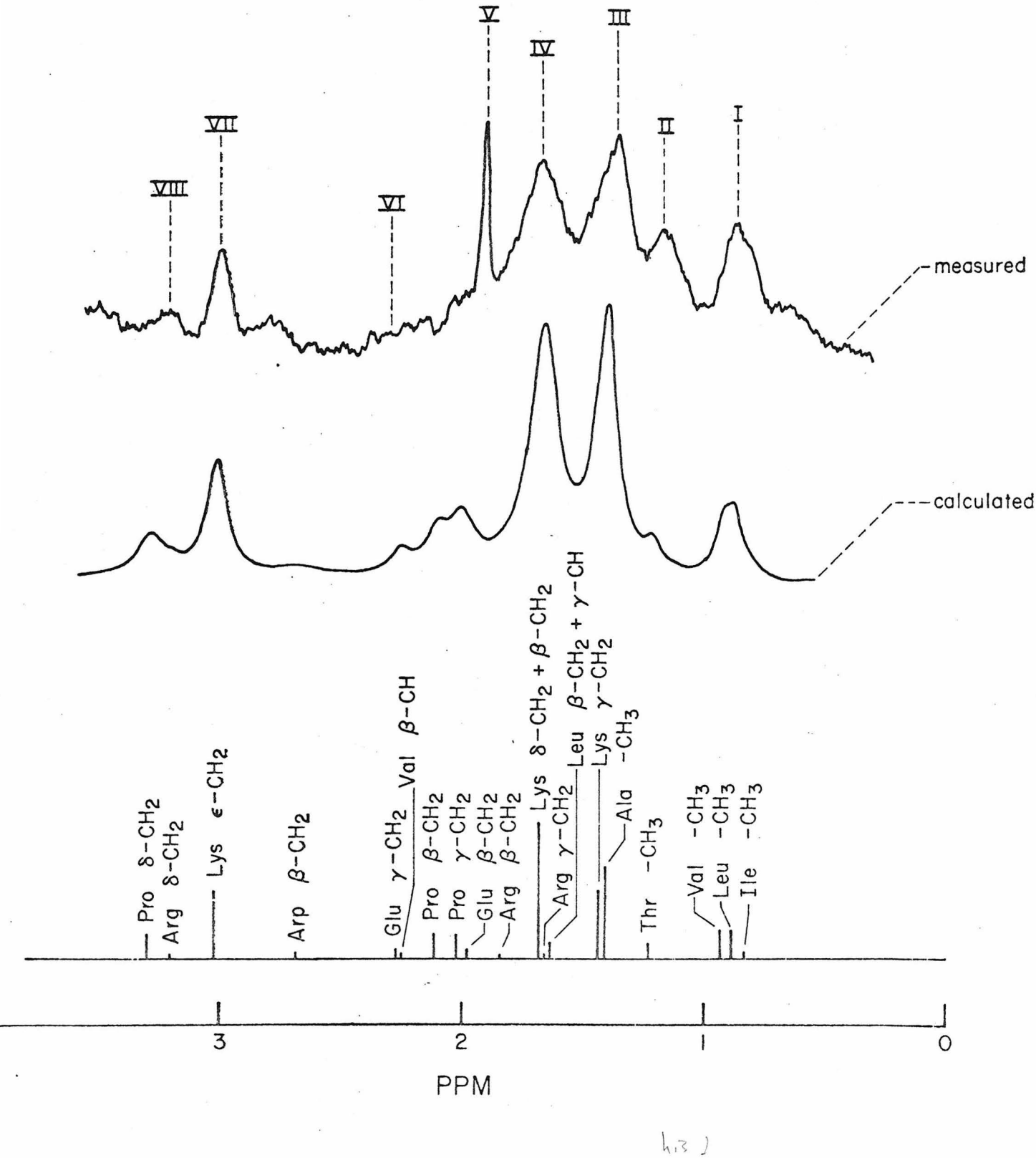
TABLE IB

Spectral regions and their corresponding proton resonances as well as percentages of the total proton concentration of histone IIb1

HIS IIb1

Spectral Region	Resonances Observed	% of Total Proton Concentration
I	Ile - CH_3 Leu - CH_3 Val - CH_3	30%
II	Ile - CH_2 Thr - CH_3	6%
III + IV	Leu - βCH_2, γCH_2 Ile - CH_2 Ala - CH_3 Lys - βCH_2, γCH_2, ϵCH_2 Arg - βCH_2, γCH_2	32%
V	Ile - β - CH Glu - β - CH_2 Pro - β - CH_2, γCH_2	12%
VI	Val - βCH_2 Glu - γCH_2 Asp - βCH_2	7%
VII + VIII	Tyr - βCH_2 Lys - ϵCH_2 His - βCH_2 Arg - δCH_2 Pro - δCH_2	13%

Figure 1. 220 MHz histone I spectrum in the high field region at 17°C and at concentration of 3 mg/ml. (The computer simulated spectrum is given at the bottom).


VIII
VII
VI
V
IV
III
II
I
measured
calculated
Pro δ-CH2
Arg δ-CH2
Lys ε-CH2
Arp β-CH2
Glu γ-CH2
Val β-CH
Pro β-CH2
Pro γ-CH2
Glu β-CH2
Arg β-CH2
Lys δ-CH2 + β-CH2
Arg γ-CH2
Leu β-CH2 + γ-CH
Lys γ-CH2
Ala -CH3
Thr -CH3
Val -CH3
Leu -CH3
Ile -CH3
3
2
1
0
PPM

h.3)

Figure 2. 220 MHz histone IIb1 spectrum in the high field region at 17°C and at concentration of 3 mg/ml.

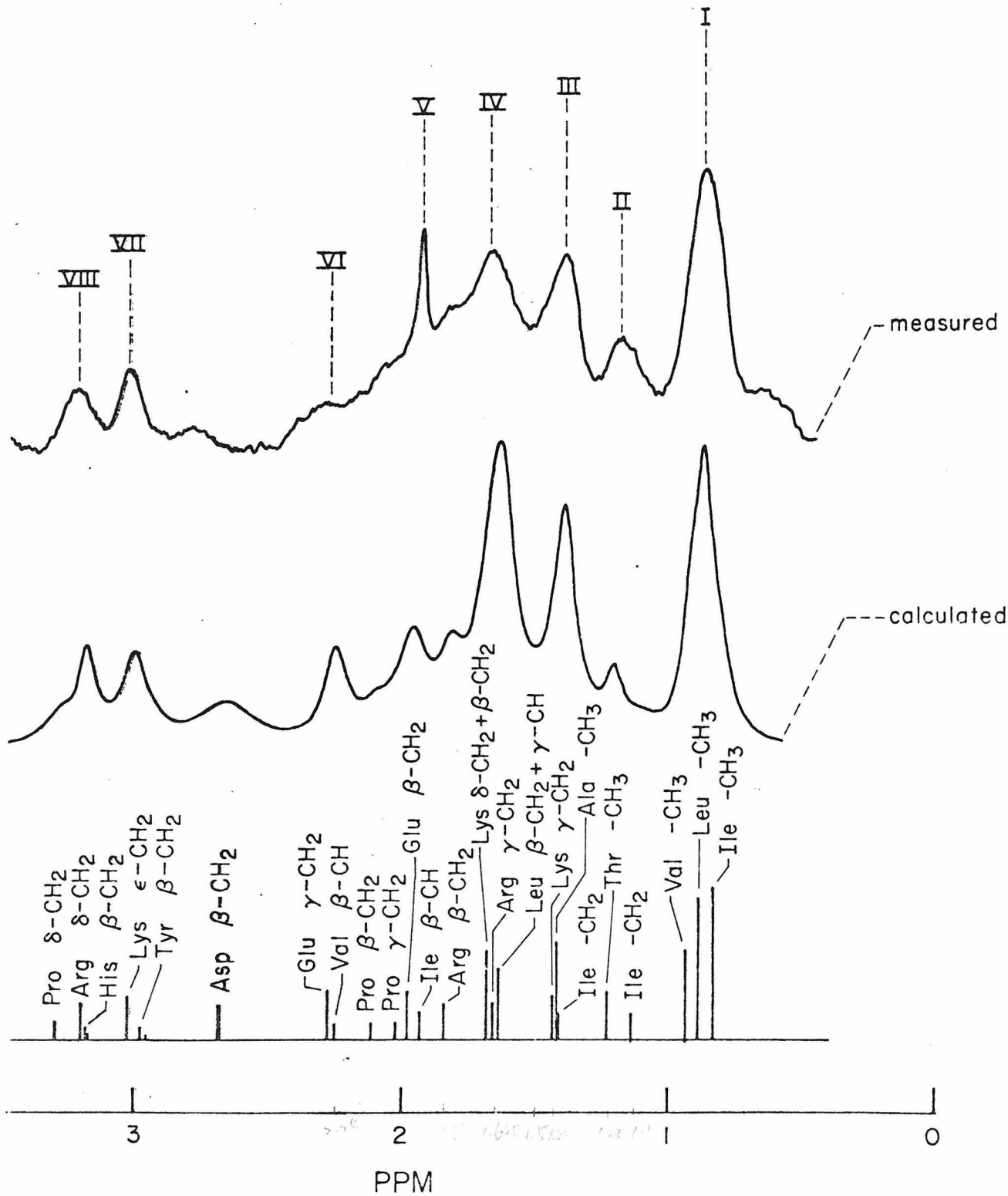
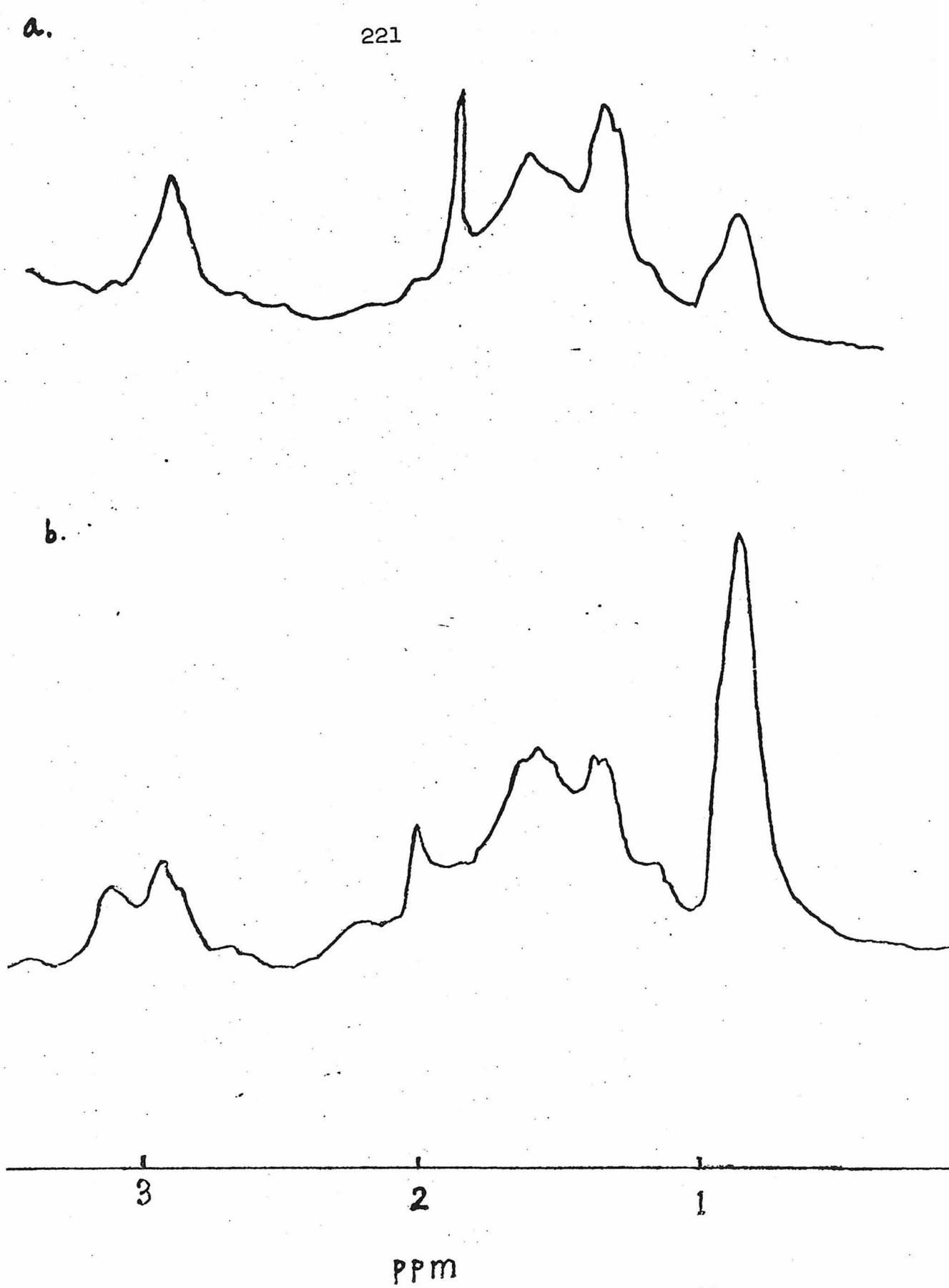

I
II
III
IV
V
VI
VII
VIII
measured
calculated
Pro δ-CH2
Arg δ-CH2
His β-CH2
Lys ε-CH2
Tyr β-CH2
Asp β-CH2
Glu γ-CH2
Val β-CH
Pro β-CH2
Pro γ-CH2
Glu β-CH2
Ile β-CH
Arg β-CH2
Lys δ-CH2+β-CH2
Arg γ-CH2
Leu β-CH2+γ-CH
Lys γ-CH2
Ala -CH3
Ile -CH2
Thr -CH3
Ile -CH2
Val -CH3
Leu -CH3
Ile -CH3
3
2
1
0
PPM

Figure 3. 100 MHz Fourier transform spectra of histone I (a) and histone IIb1 (b). Number of transients 400.


a.
b.
3
2
1
ppm

(2). Concentration Dependence

Although qualitatively the spectra of histone I and histone IIb1 do not change significantly within the concentration range 3 mg/ml to 50 mg/ml in the salt free aqueous solution, the observed spectral area does not increase proportionally with increasing protein concentration. The results of the spectral area measurements at various protein concentrations are expressed in terms of the percentage of the total proton concentrations of the proteins and are given in Table I and Figs. 4 and 5. From these, it is noted that the concentration dependent area loss is most significant for the side chain protons of the aliphatic amino acid residues (Region I) and the terminal ε-CH_2 group of lysine as well as arginine residues (Regions VII and VIII) of these proteins. At low protein concentration ($\sim$ 3 mg/ml), essentially all the expected proton intensity was observed ($\simeq$90%). Compared to the total proton concentrations expected for these proteins, the percentage of the spectral area observed is only 55% in the case of histone I and is about 65% in the case of histone IIb1, when the protein concentration is increased to 50 mg/ml.

(3). Salt Dependence

At a given protein concentration ($\sim$ 3 mg/ml), the observed area in some spectral regions decreases progressively with increasing NaCl concentration. No significant changes in chemical shifts of each spectral region were observed upon the addition of salt. In the case of histone I,

TABLE I. Percentage of the Total Proton Concentrations Observed

Histone I

Concentration	Spectral Region					
	I	II	III + IV	V	VI	VII + VIII
Co[a]	98	99	90	99	99	99
2 Co	93	99	85	95	95	90
4 Co	85	75	75	80	80	80
8 Co	73	70	70	60	60	72
16 Co	63	70	65	45	40	60

Histone IIb1

Co	90	99	98	90	95	98
2 Co	81	93	97	88	90	95
4 Co	77	85	95	85	80	88
8 Co	73	82	92	72	65	80
12 Co	66	80	86	65	50	70
16 Co	61	78	82	55	45	62

[a] Co = 3.1 mg/ml

Figure 4. Concentration dependence of histone I spectra.

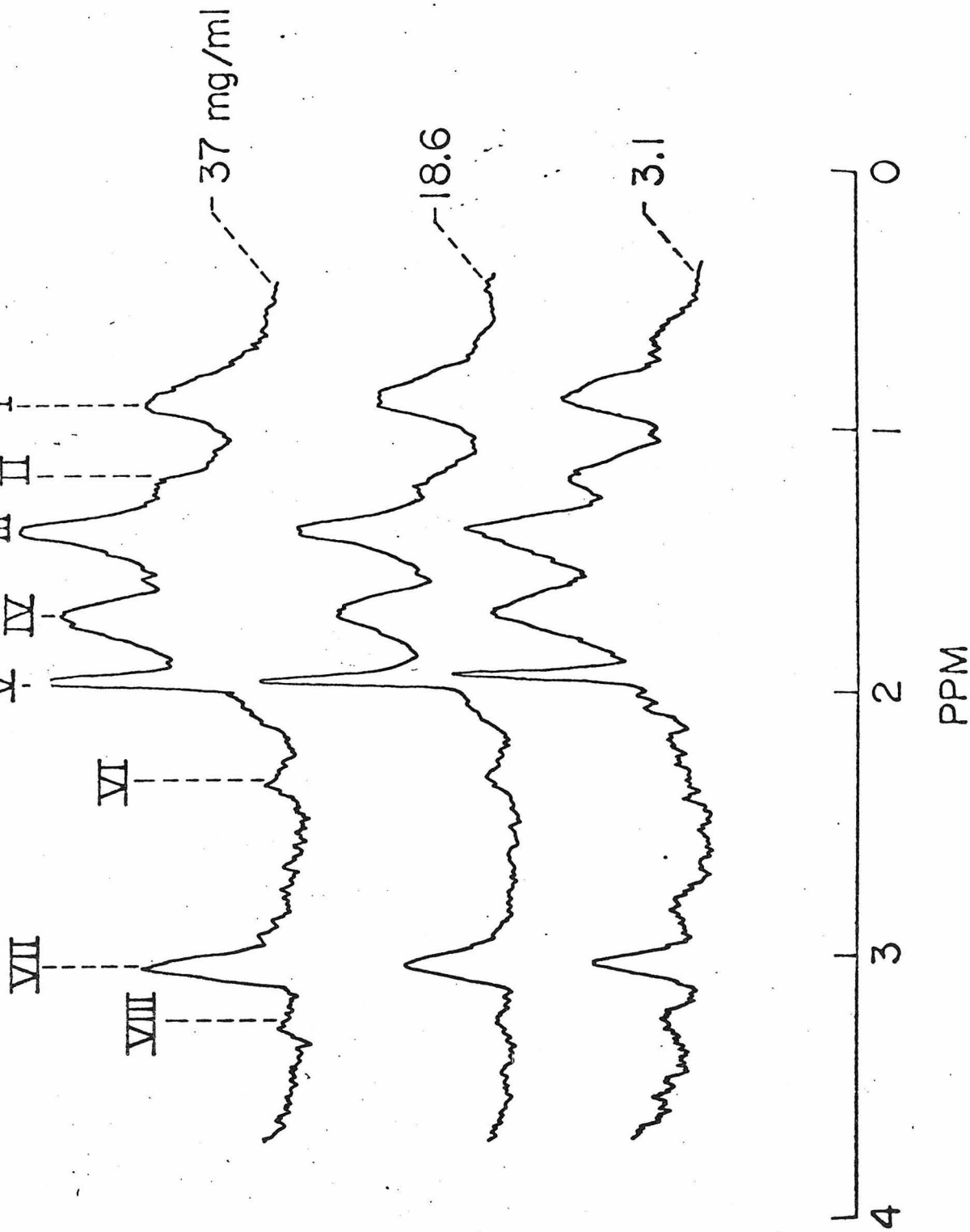
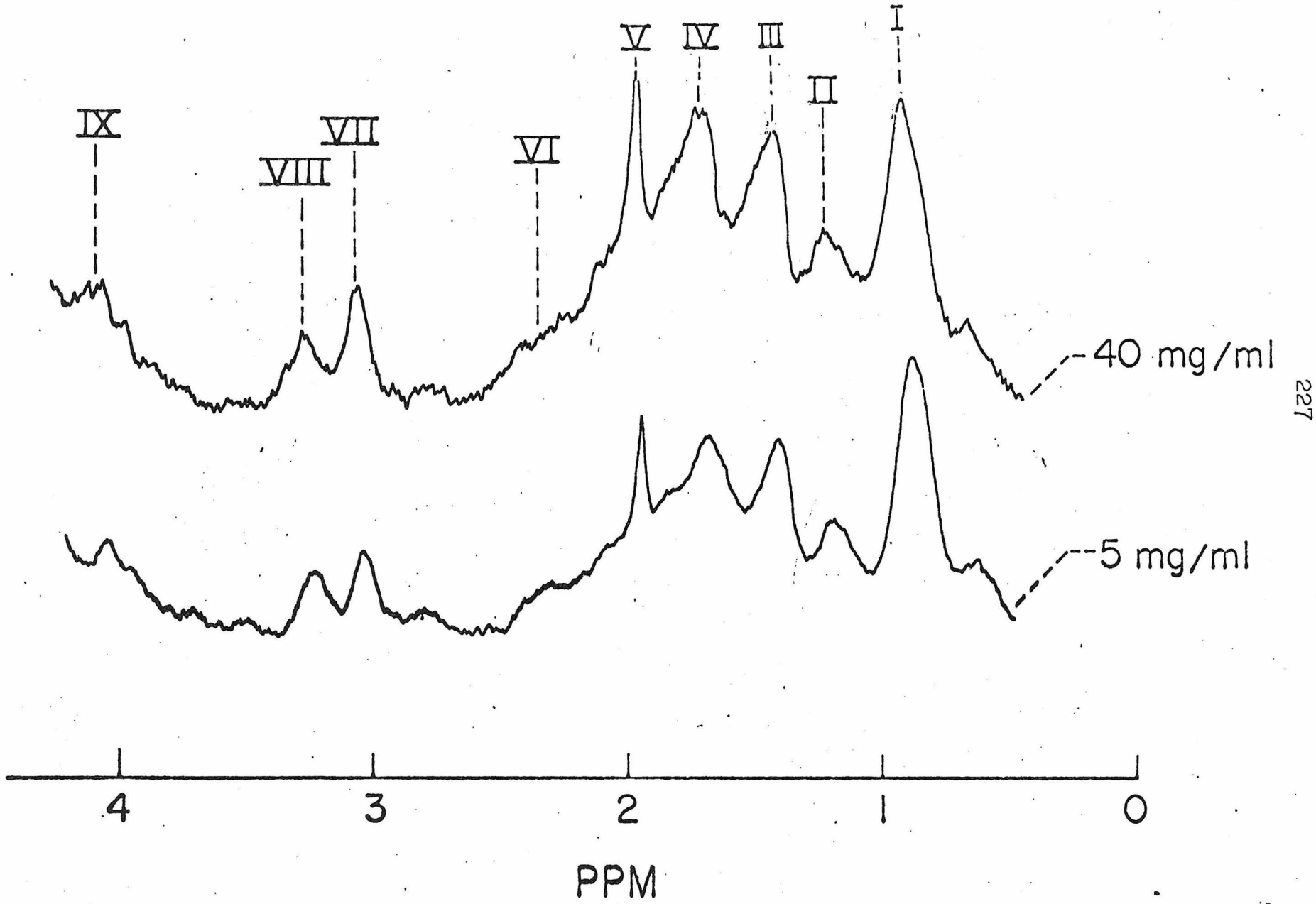

37 mg/ml
18.6
3.1
I
II
III
IV
V
VI
VII
VIII
4
3
2
1
0
PPM

Figure 5. Concentration dependence of histone IIb1 spectra.


IX
VIII
VII
VI
V
IV
III
II
I
40 mg/ml
5 mg/ml
4
3
2
1
0
PPM

the salt induced area loss of the spectral region I is more preferential than any other spectral region. While in histone IIb1, an equally significant area loss was observed for each spectral region upon the addition of salt. The high field spectra of histone I and histone IIb1 under different NaCl concentrations are shown in Fig. 6 and Fig. 7. The concentration study of histone I in the presence of 0.5 M NaCl exhibited a concentration dependence of the spectral area loss for all the spectral regions especially for region I where only 45% of the spectral area was observed compared to that expected for the total proton concentration of the protein at 25 mg/ml. (Table II) . In the case of histone IIb1, the concentration study in the presence of 0.15 M NaCl also indicated that the percentage of the spectral area observed decreases with increasing protein concentration; especially for Region I and VII, where only about 50% of spectral area was observed at the protein concentration of 25 mg/ml.

The addition of other salts such as $MgCl_2$, $NaClO_4$ also shows a similar effect on the spectra of both proteins. For example in the case of histone I at concentration 3 mg/ml, 50% of spectral area in Region I was lost in the presence of either 0.4 M $MgCl_2$ or 0.2 M $NaClO_4$. In histone IIb1, 80% of the spectral area in Region I was lost in the presence of either 1 M NaCl, 0.3 M $MgCl_2$, or 0.1 M $NaClO_4$. The percentages of the total spectral area observed at different salt concentrations are given in Tables III and IV and Fig. 8.

Figure 6. NaCl dependence of histone I spectra.

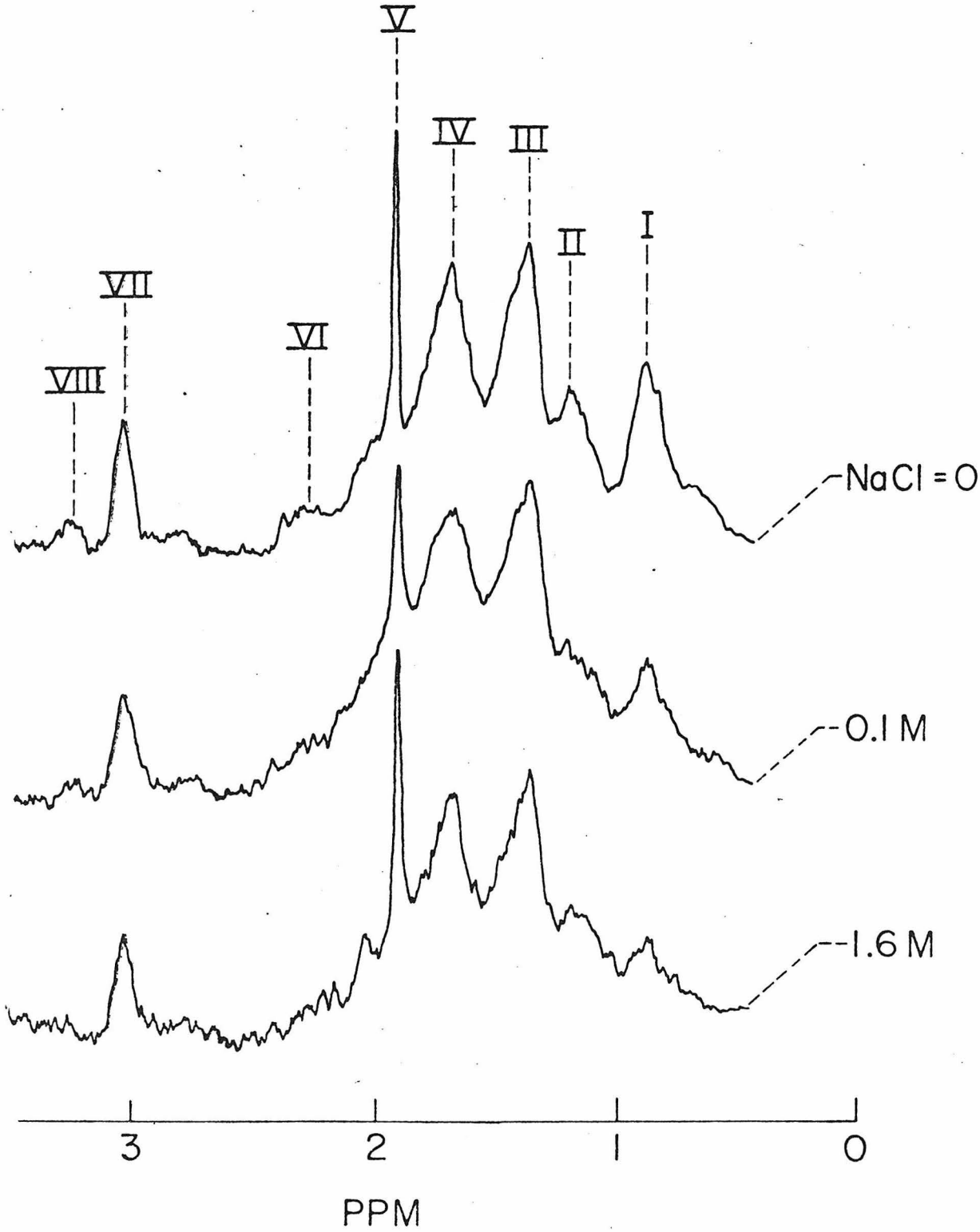

V
IV
III
II
I
VII
VI
VIII
NaCl = 0
0.1 M
1.6 M
3
2
1
0
PPM

Figure 7. NaCl dependence of histone IIb1 spectra.

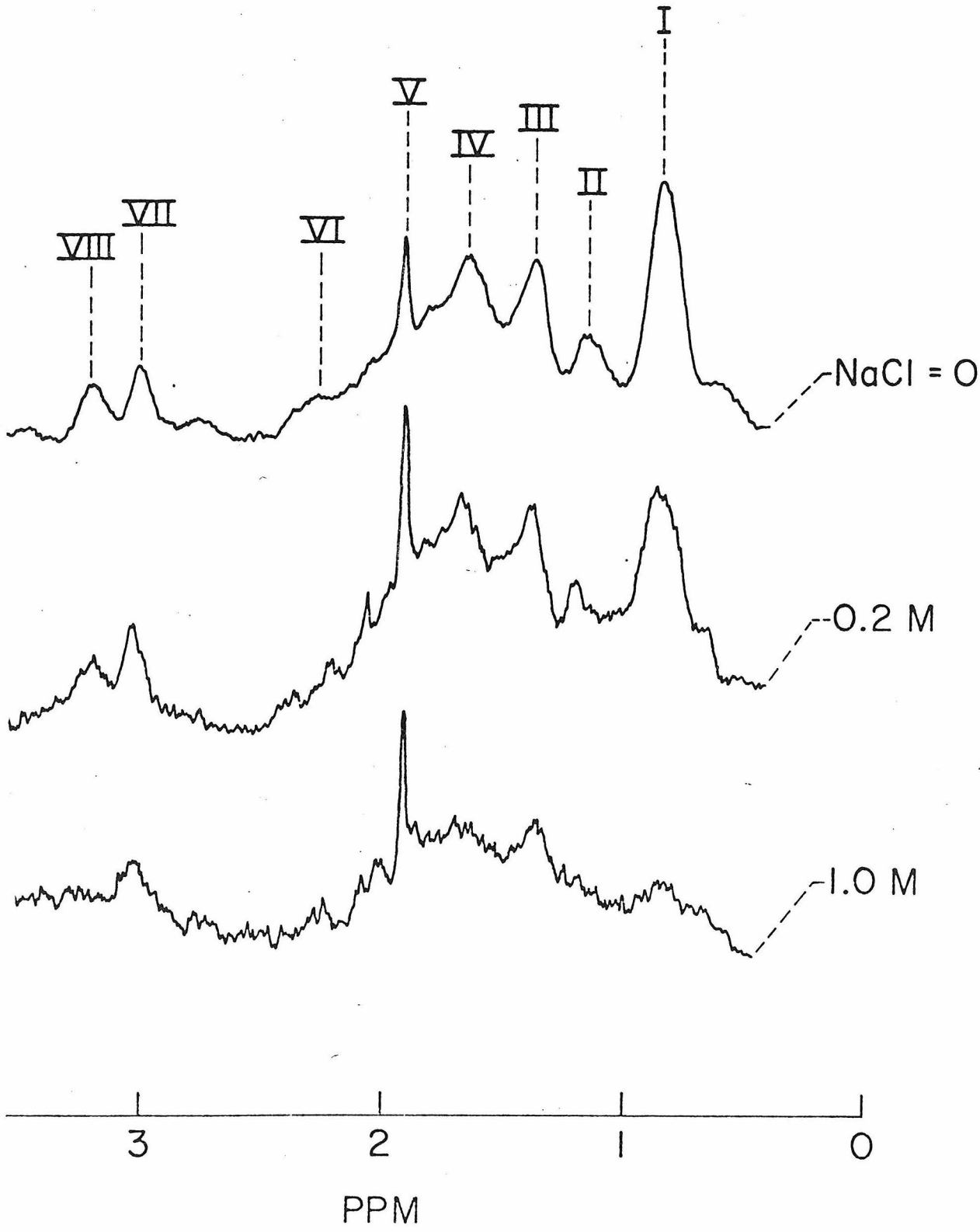

I
II
III
IV
V
VI
VII
VIII
NaCl = O
0.2 M
1.0 M
3
2
1
0
PPM

Figure 8. Dependence of the spectral area of Region I of histones on the salt concentrations.

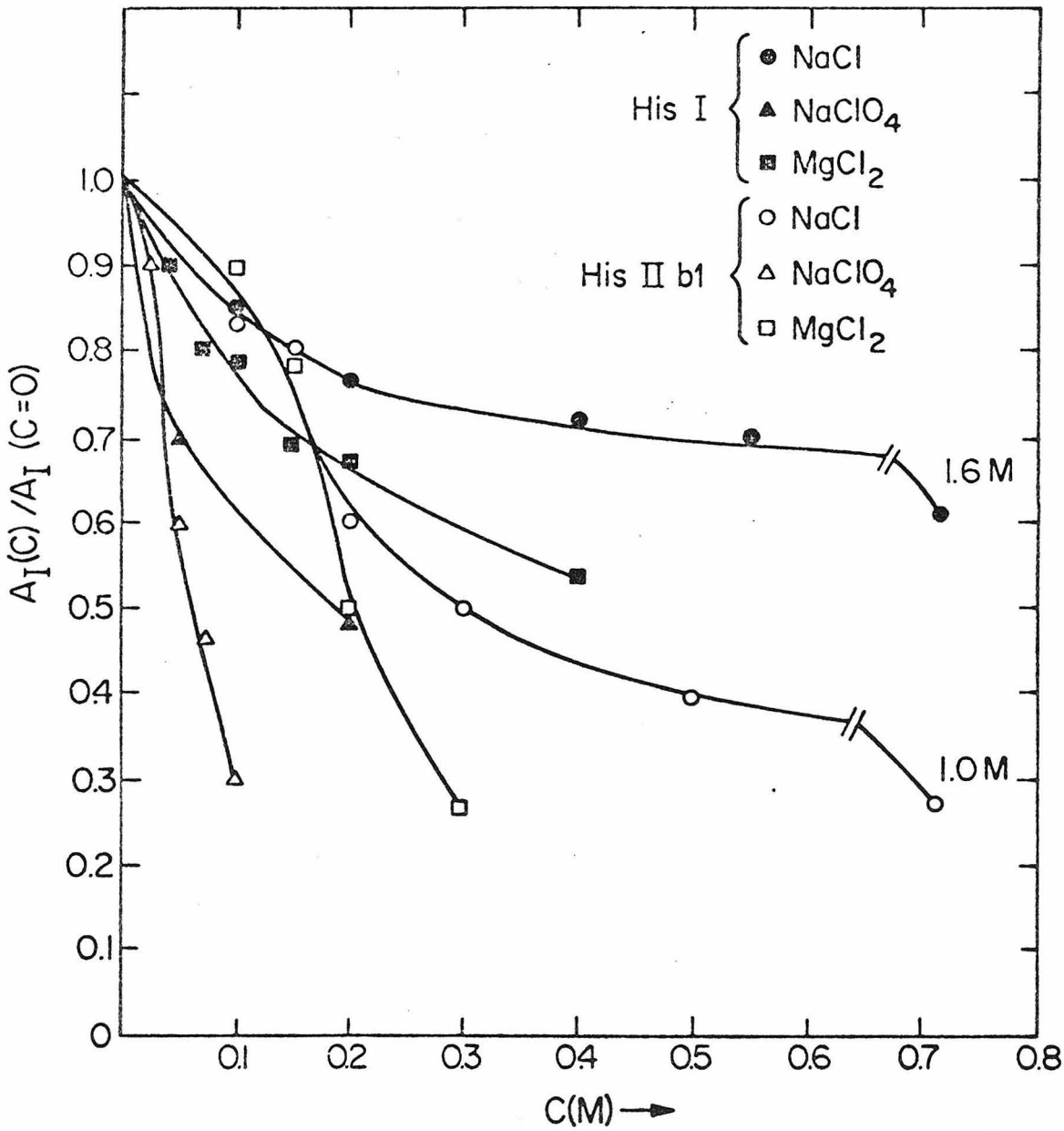

His I
NaCl
NaClO4
MgCl2
His II b1
NaCl
NaClO4
MgCl2
1.6 M
1.0 M
AI(C) /AI (C=0)
C(M)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

TABLE II. Percentage of the Total Proton Concentrations Observed

Histone I (0.5 M NaCl)

Concentration	Spectral Region					
	I	II	III + IV	V	VI	VII + VIII
Co[a]	85	90	90	95	90	90
2 Co	70	85	75	90	85	75
4 Co	55	85	64	80	80	68
8 Co	45	85	58	70	70	55

Histone IIb1 (0.15 M NaCl)

2 Co	70	80	80	90	90	80
4 Co	60	80	75	81	80	70
8 Co	52	78	70	65	65	60
12 Co	48	72	62	50	55	54

[a] Co ≈ 3.1 mg/ml

TABLE III Percentage of the Total Proton Concentration Observed

Histone I

[NaCl]	Spectral Region I	VII	A_{VII}/A_I
0.1 M	82	85	0.34
0.2 M	71	76	0.41
0.4 M	65	70	0.41
0.8 M	50	65	0.49
1.6 M	35	55	0.74
[$NaClO_4$]			
0.05 M	55	70	0.44
0.10 M	50	55	0.30
0.20 M	45	35	0.25

TABLE IV. Percentage of the Total Proton Concentration Observed

Histone IIb1

Spectral Region

	I	II	III + IV	VII + VIII
[NaCl]				
0	85	93	97	95
0.1 M	75	75	85	83
0.2 M	60	70	71	54
0.3 M	45	60	61	40
0.5 M	40	45	48	
1.0 M	25		45	35
[$NaClO_4$]				
0.025 M	75		90	85
0.050 M	51		80	60
0.075 M	40		63	45
0.100 M	35		55	35
[$MgCl_2$]				
0.05 M	80		85	90
0.075 M	68		80	80
0.100 M	55		65	75
0.150 M	40		45	50

(4). pD Dependence

As shown in Fig. 9 and Fig. 10, the pmr spectra of both histone I and histone IIb1 are pD dependent, especially at pD < 1.5 and pD > 10.0. In the case of histone I, the observed area of each spectral region is essentially unchanged at neutral pD's at a given protein concentration. However, starting at pD ∿ 9.0, the observed spectral area of Region I as well as other spectral regions decreases with increasing solution pD until pD ∿ 11.0, after which the spectral area increases abruptly with increasing pD. The spectral area observed at pD ∿ 11.0 in Region I is only 40% of that observed at pD 7.0. A significant high field shift of Region V from pD 4.0 to 6.0 indicates the deprotonation of the γ carboxyl group of glutamic acid residues of the protein. Similarly as pD increases from 9.0 to 11.0, the deprotonation of terminal NH_3^+ group of lysine residues also induces about 100 Hz of high field shift to spectral region VII corresponding to ε-CH_2 group of these residues. (Fig. 11).

In the case of histone IIb1, a similar phenomenon was also observed, but the area loss around pD 11.0 is even more pronounced, where almost 90% of the spectral area in Region I is lost. (Fig. 12).

(5). Temperature Dependence

At high protein concentrations (∿25 mg/ml) in the absence of salt, the observed spectral area of each spectral region of the proteins increases with increasing temperature from 17°C to 55°C, especially for

Figure 9. pD dependence of histone I spectra.

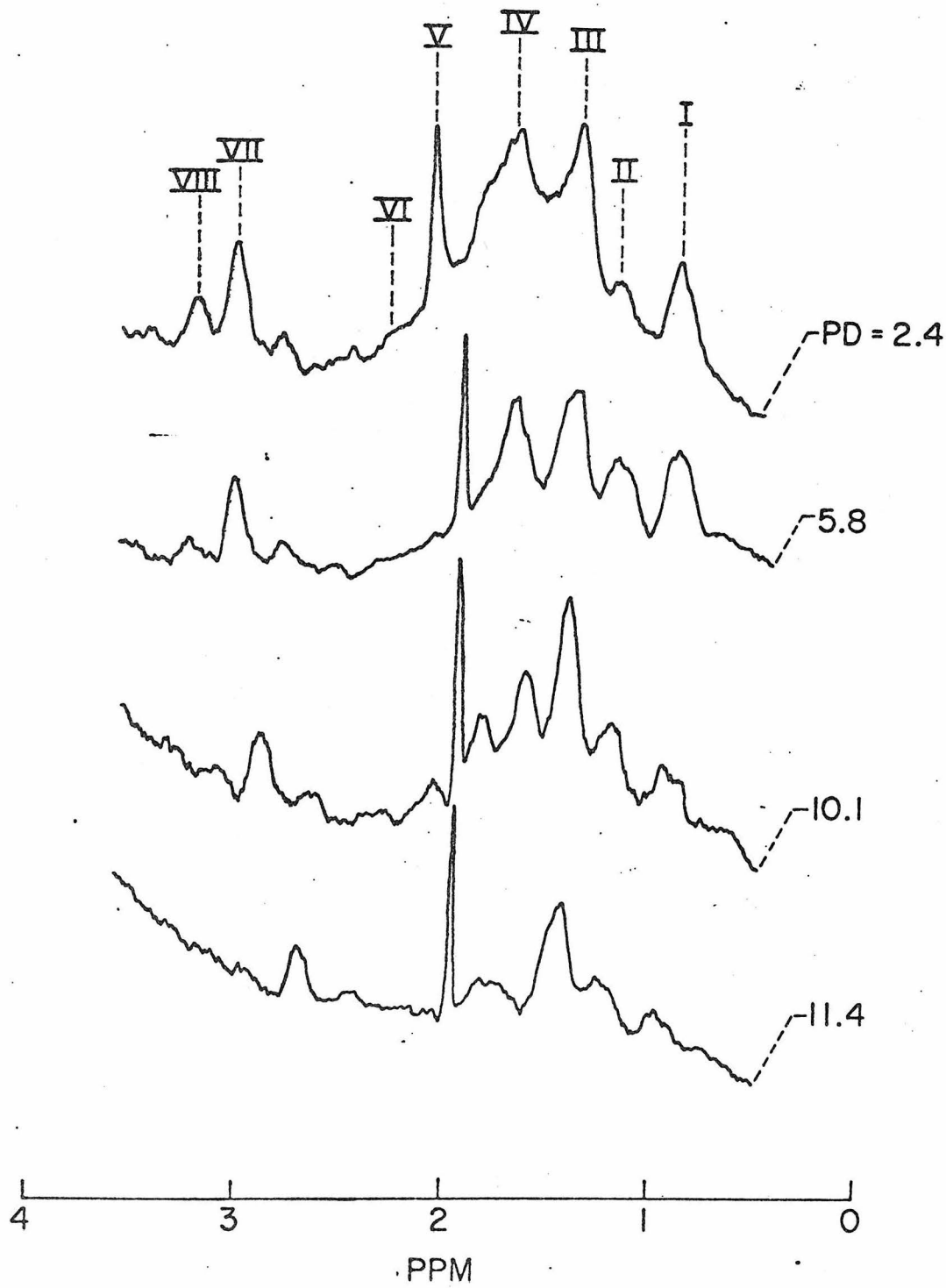

V
IV
III
I
VIII
VII
VI
II
PD = 2.4
5.8
10.1
11.4
4
3
2
1
0
PPM

Figure 10. pD dependence of histone IIb1 spectra.

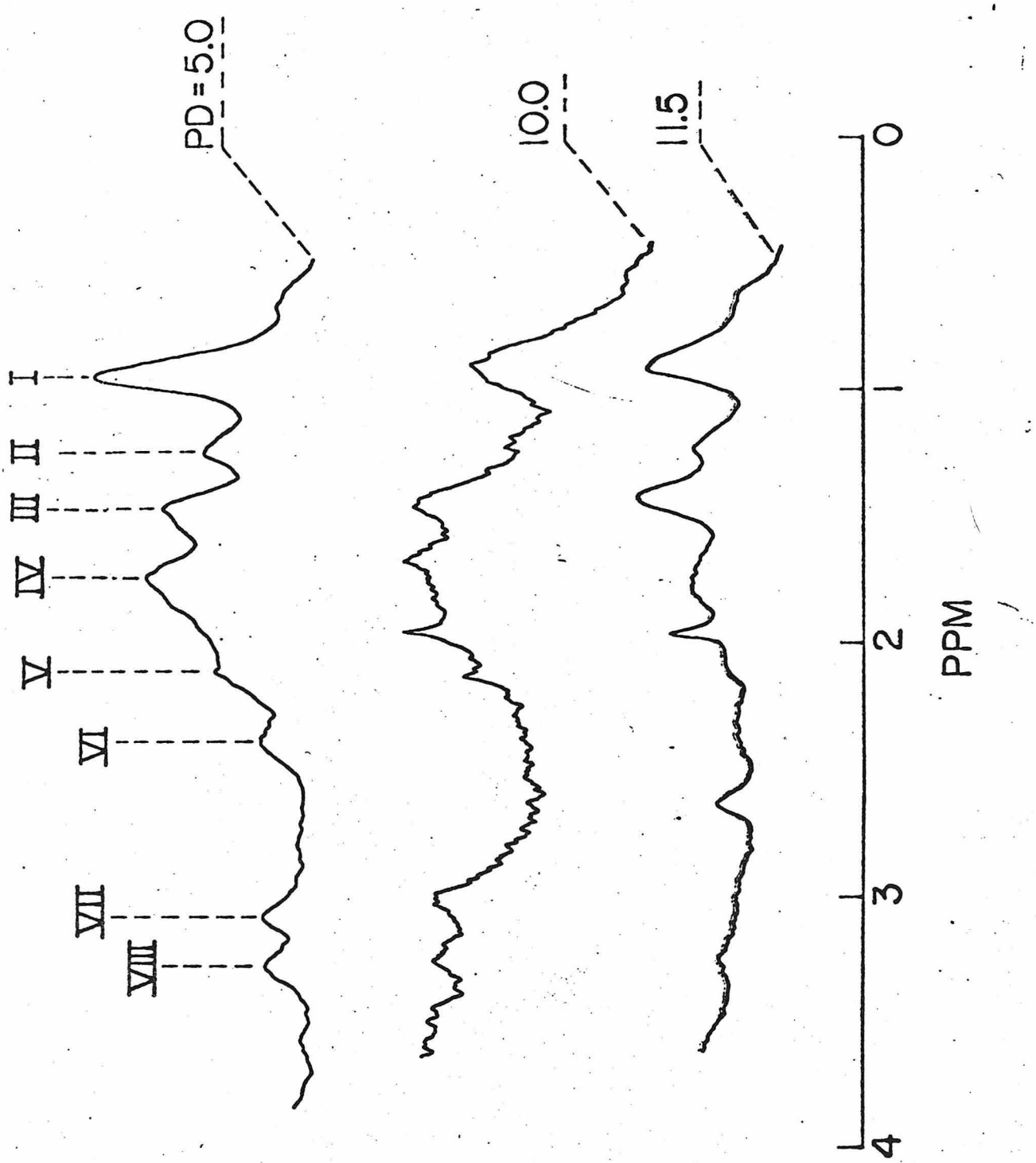

PD=5.0
10.0
11.5
I
II
III
IV
V
VI
VII
VIII
0
1
2
3
4
PPM

Figure 11. Dependence of the spectral area of Region I of histone I on the solution pD (→ indicates the direction of titration).

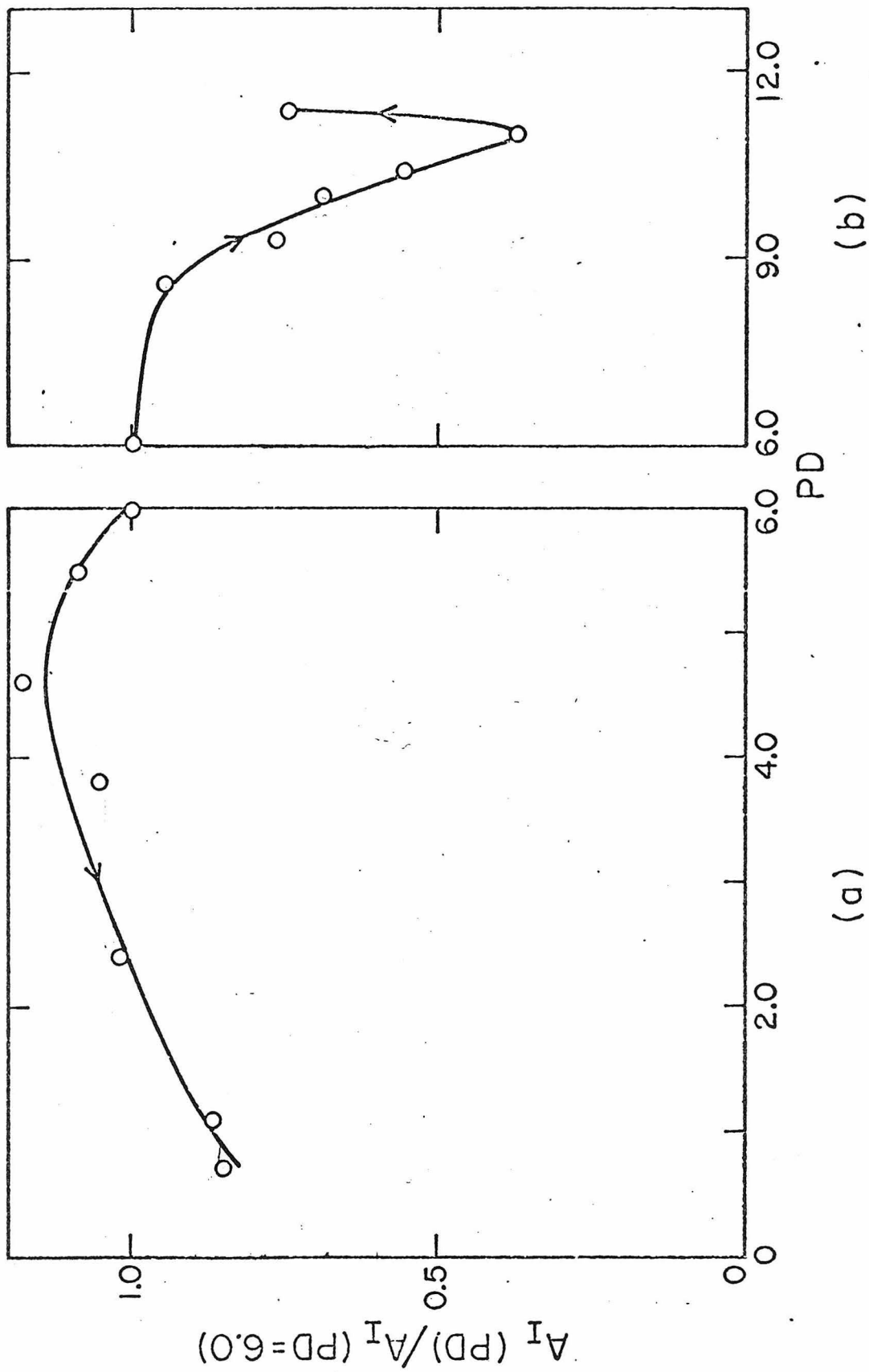
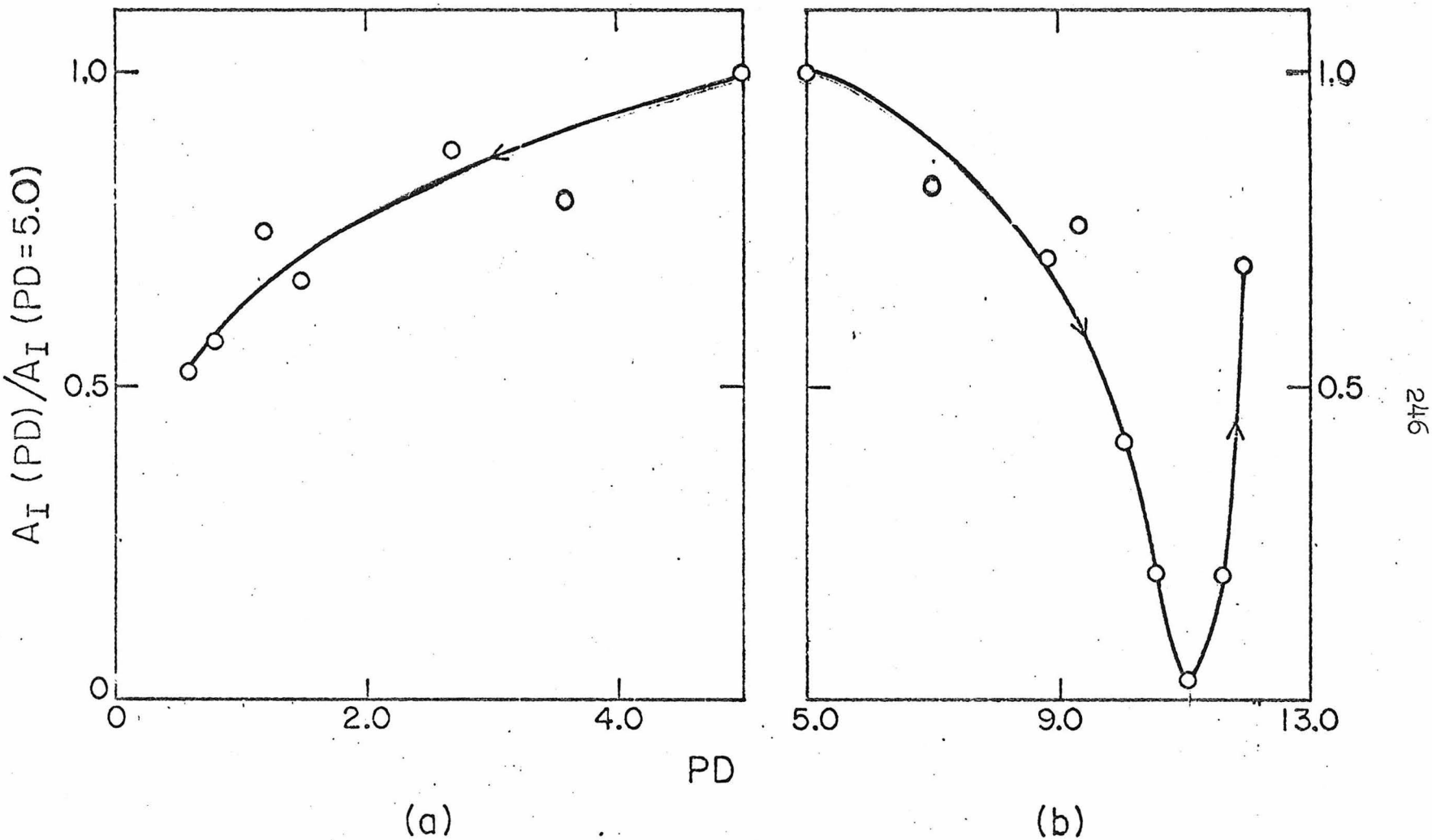

A_I (PD)/A_I (PD=6.0)
0
0.5
1.0
0
2.0
4.0
6.0
(a)
6.0
9.0
12.0
(b)
PD

Figure 12. Dependence of the spectral area of Region I of histone IIb1 on the solution pD (→ indicates the direction of titration).


1.0
0.5
0
A_I (PD)/A_I (PD=5.0)
0
2.0
4.0
PD
(a)
1.0
0.5
5.0
9.0
13.0
(b)

Region I corresponding to the side chain protons of the aliphatic amino acid residues. For example in histone I at 25 mg/ml, the observed percentage of the spectral area in Region I increases from 70% at 17°C to 85% at 55°C. In histone IIb1 at the same concentration, the observed area of this region also increases from 66% at 17°C to 87% at 35°C. Similar increases of the spectral area in other regions with increasing temperature were also observed.

In the presence of 0.5 M NaCl, the observed spectral area of Region I increases from 70% at 17°C to 85% at 35°C for histone I at the concentration of 6 mg/ml. While in the presence of 0.3 M NaCl, the area observed in the spectral region I of the same concentration of histone IIb1 increases gradually from 45% at 17°C to 65% at 55°C.

4. DISCUSSION

At low protein concentrations, histone I and histone IIb1 are essentially random coil polypeptides in neutral salt free aqueous solution with well defined spectral regions corresponding to various side chain protons of the amino acid residues of these proteins. This conclusion was confirmed by the comparison of the calculated spectra based on the model of random coil polypeptides and was also supported by the results of ORD and CD studies.[4]

However, a significant area loss in some spectral regions of these proteins was observed either at high protein concentrations, in the

presence of salt or at extreme pD's where the extensive protein aggregations are known to occur. In order to understand how this spectral area loss is related to the protein aggregations, we should know first the origin and the mechanisms of this phenomenon from the known theory of nmr.

(1). Mechanism of Spectral Area Loss[13]

In nmr spectroscopy, it is well known that the transverse relaxation time (T_2) of a given resonance in a molecular system can properly be represented by its half-width, if the contribution from T_2 exceeds that from the field inhomogeneity. In case that motional freedom is reduced because of the molecular aggregations, the dipole-dipole interactions arising from the close contact or the rigidity of the environment in the aggregated species will significantly shorten the transverse relaxation times of the resonances of the aggregated species. For example, in the random coil state, the resonances of the amino acid side chain protons of nucleoproteins normally exhibit 10 to 30 Hz of halfwidth in the high resolution pmr spectra at room temperatures. However upon molecular aggregations, the resonances of some amino acid side chain protons in the contact or rigid region of the aggregated species will broaden to several hundred Hz of the spectral width. The appearance of the resulting pmr spectra in a system of extensive molecular aggregations is sensitive to the association rates occurring on the nmr time scale of observation. This "critical rate" can be thought

of as the inverse of the characteristic time in which the transverse components of the magnetization dephase after exchange from one site to another. Usually this "critical rate" is determined by the chemical shift differences (Δ) between the two sites (see Part I, Section 3.4), provided that the transverse relaxation times are long compared to the inverse of Δ and comparable in these two sites. However in the case of histone aggregations, the "critical rate" is determined by the inverse of the transverse relaxation times of the resonances in the aggregated species, since the relaxation times of the aggregated species are much shorter than those of the unaggregated random coiled histone molecules. If the rates of histone associations are faster than this "critical rate", the weight-averaged spectra will appear. The halfwidths of the resonances in the observed spectra would be the weight average of those in the aggregated and unaggregated species. But this is not the case observed in our system. On the other hand, if the association rate is very slow compared to this "critical rate", the resonances arising from the aggregated and unaggregated species are separated. Under this condition, only the resonances of the unaggregated species can be observed in the pmr spectra, because the resonances of the aggregated species are supposedly too broad to be observed. As a result, the spectral area loss is expected to occur in such a system of "slow" protein aggregations.

Histone I and histone IIb1 are known to have low α-helix content (10 to 20%) which is independent of the protein concentrations and is slightly salt dependent. A calculation of the transverse relaxation

time of histone molecules based on the transition from a random coil to full α-helix transition showed that it should produce a doubling of the spectral width but no significant area loss. The progressive loss of spectral area upon increasing the protein or salt concentrations indicates that it does not arise from random coil to α-helix transition of these proteins but from other molecular aggregation processes. A strong concentration dependence of the spectral area loss also indicates that it mainly arises from the processes of intermolecular aggregations.

The fact that the area loss was observed in all the spectral regions indicates that all the amino acid residues of these proteins are involved in the protein aggregations and both the hydrophobic and electrostatic interactions are essential for these aggregation processes. The area loss in Region I would indicate that there are intimate hydrophobic contacts among the side chains of the aliphatic amino acid residues in the aggregated species. On the other hand, the area loss in the spectral regions corresponding to the terminal side chain protons of basic as well as acidic amino acid residues suggests the significance of the electrostatic interactions between the terminal anionic carboxyl group of glutamic or aspartic acid residues and the terminal ammonium group of lysine residues or guanidine group of arginine residues during the protein aggregations.

The phenomenon of the spectral area loss is even more pronounced when the salt is present in the solution. In the case of histone I at low concentrations, a preferential area loss in the spectral region I

upon the addition of NaCl indicates that the hydrophobic interactions become more favorable for the protein aggregations in the presence of salt. (Table III). This is probably due to the fact that, in the presence of salt, the increase in counter ion condensation to the charged amino acid residues of this protein would shield considerably the long range electrostatic interactions and enhance the hydrophobic interactions for the protein aggregations. While in histone IIb1, all the spectral regions experience almost the same extent of area loss upon the addition of NaCl. (Table IV). This would indicate that both the electrostatic and hydrophobic interactions are involved in the protein aggregation in the presence of high concentration of NaCl. The difference in spectral behavior in the presence of salt could be attributed to the difference in the ratios of the basic to acidic amino acid residues and their distributions in these two proteins.

(2). Model Study of Histone Aggregations

In order to understand the quantitative relations between the spectral area loss and the degree of molecular associations, the concentration study of these proteins was carried out either in the presence or in the absence of salt.

From the concentration dependence of the spectral area loss of histone I and histone IIb1, it is known that there are intermolecular aggregations in aqueous solution upon increasing protein concentrations. Several models are therefore proposed in the following in order to fit

our experimental observations. All of these models are simply based on the assumption that the exchange rate for the protein aggregation processes is slow compared to the inverse of the transverse relaxation times of the aggregated species and the resonances observed in the high resolution pmr spectra are primarily due to the unaggregated species.

Model 1. (n-mer model)

Assuming that the protein aggregation is due to one-step process, we can use the following equation to describe this process.

$$n \cdot H \underset{}{\overset{K_n}{\rightleftharpoons}} H_n$$

where H represents the unaggregated monomer of the protein, H_n is the aggregated species due to the association of the n monomers of the protein, n is the number to be determined, K_n is the equilibrium constant.

According to our assumption, the observed area of certain spectral region A_{obs} would be proportional to the concentration of the unaggregated species in the solution [H].

$$\therefore \quad A_{obs} = k\,[H]$$
$$A_o = k[H_o] \quad ,$$

where k is proportionality constant, A_o represents the spectral area corresponding to the original total histone concentration $[H_o]$ without

any aggregations

$$K = \frac{[H_n]}{[H]^n} \quad (1) \qquad [H_o] = [H] + n[H_n] \quad (2)$$

$$[H_n] = \frac{[H_o] - [H]}{n} \quad (3)$$

Substitute (2) and (3) into eq. (1) and take the log on both sides of eq. (1). We get $\log [[H_o] - [H]) = n \log [H] + \log nK$ or $\log (A_o - A_{obs}) = n \log A_{obs} + \log nK + (n - 1) \log k$. From this plot, we may get a straight line of either from $\log ([H_o] - [H])$ vs. $\log [H]$ or from $\log (A_o - A_{obs})$ vs. $\log A_{obs}$ with slope n. (Fig. 13). The equilibrium constant can also be obtained from this plot. In some spectral region where the area loss is incomplete upon protein aggregations, there are still some resonances in the aggregated species contributing to the observed spectral area. In this case a correction factor should be added to the expressions of the observed spectral area, i.e.,

$$A_{obs} = k([H] + n[H_n] \cdot f) = k[H]_{obs}$$

where f is the fraction of the resonances of the aggregated species which are observed in the spectra.

$$\therefore K = \frac{(1 - f)^{n-1} \quad ([H_o] - [H_{obs}])}{n \cdot ([H]_{obs} - f[H_o])^n} \quad (4)$$

After taking the log on both sides of the eq. (4) we obtain the following equation.

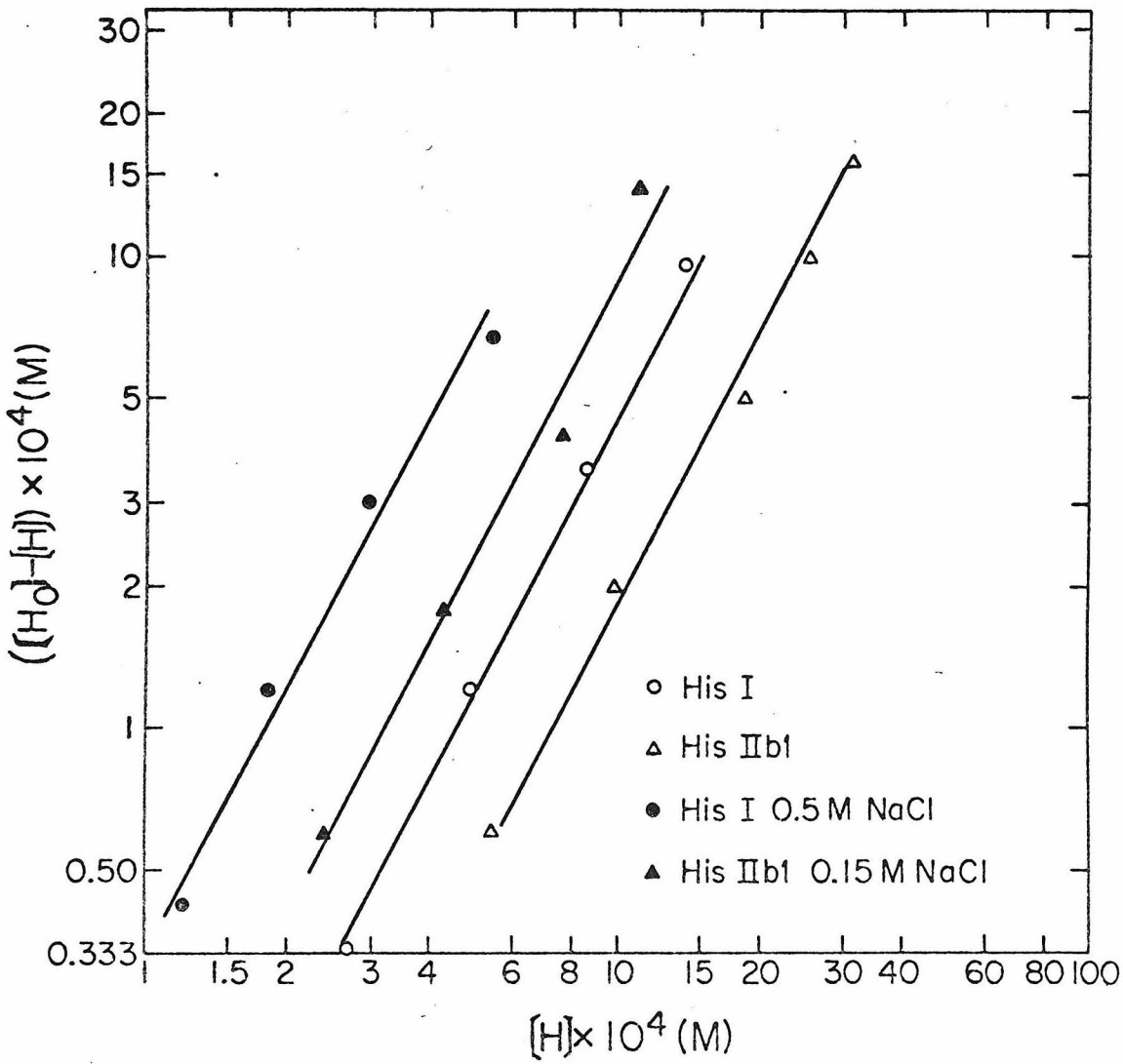
$$\log ([H_o] - [H]_{obs}) = n \log ([H]_{obs} - f[H]_o)$$

$$+ \log n K - (n - 1) \log (1 - f) \quad .$$

From the plot of $\log ([H_o] - [H]_{obs})$ vs. $\log ([H]_{obs} - f[H]_o)$ we can also get a straignt line with slope n, if this model is adequate to describe the histone aggregations.

In the case of either histone I or histone IIb1 in the salt free solution, n is obtained as $\sim$ 2.0 for both the spectral region I and VII + VIII, with f = 0, and also for the spectral region III + IV with f = 0.1. This would indicate that the area loss is not complete due to protein aggregation for the spectral regions corresponding to the middle of the side chain protons of lysine and arginine residues. In the presence of 0.5 M NaCl, histone I also exhibits slope n $\sim$ 2.0 in this plot for the spectral regions, I, III + IV and VII + VIII. The same result was also obtained for histone IIb1 in the presence of 0.15 M NaCl. This model indicates that upon histone aggregations, the dimer is the most favorable aggregated species being formed.

Figure 13. Double logarithm plots of $([H_o] - [H]) \times 10^4$ vs. $[H] \times 10^4$, where $[H^o]$ is the stoichiometric concentration of histone and [H] is histone in the free state. (Model 1).


30
20
15
10
5
3
2
1
0.50
0.333
1
1.5
2
3
4
6
8
10
15
20
30
40
60
80
100
([H]o-[H]) x 10^4 (M)
[H] x 10^4 (M)
His I
His IIb1
His I 0.5 M NaCl
His IIb1 0.15 M NaCl

Model 2. (polymeric model with the same equilibrium constant for each step)

Consider the following successive equilibrium steps upon protein aggregation:

$$H_1 \overset{K_2}{\rightleftharpoons} H_2 \overset{K_3}{\rightleftharpoons} H_3 \overset{K_4}{\rightleftharpoons} H_4 \rightleftharpoons \cdots\cdots \overset{K_n}{\rightleftharpoons} H_n$$

where H_1, H_2 $\cdots$ H_n denote the monomer, dimer and n-mer of the protein molecules. $K_2 = K_3 = K_4 \cdots = K_n = K$ is equilibrium constants for each step. Assuming that only the monomers can be observed in the pmr spectra, the total histone concentration $[H_o]$ can be expressed in terms of the monomer concentration, i.e.,

$$[H_o] = [H_1] + 2[H_2] + \cdots\cdots + n[H_n]$$

$$[H]_{obs} = [H_1] \ , \ K_2 = K_3 \cdots = K_n = K$$

$$[H_o] = [H_1]\{1 + 2K_2[H_1] + n \cdot K_2K_3 \cdots K_n [H]^{n-1}\}$$

$$= [H_1] \sum_{n=1}^{\infty} n(K[H_1])^{n-1}, \text{ if } K[H_1] << 1$$

$$[H_o] = \frac{[H_1]}{(1 - K[H_1])^2} = \frac{[H]_{obs}}{(1 - K[H]_{obs})^2}$$

$$X = \frac{[H]_{obs}}{[H_o]}$$

$$\frac{1}{X} - \frac{1}{\sqrt{X}} = K[H_o]$$

Figure 14. Plot of $1/x - 1/\sqrt{x}$ vs. $[H_o] \times 10^4$ (Model 2, histone I) where x is the fraction of histone I in thr free state and $[H_o]$ is the stoichiometric concentration of histone I.

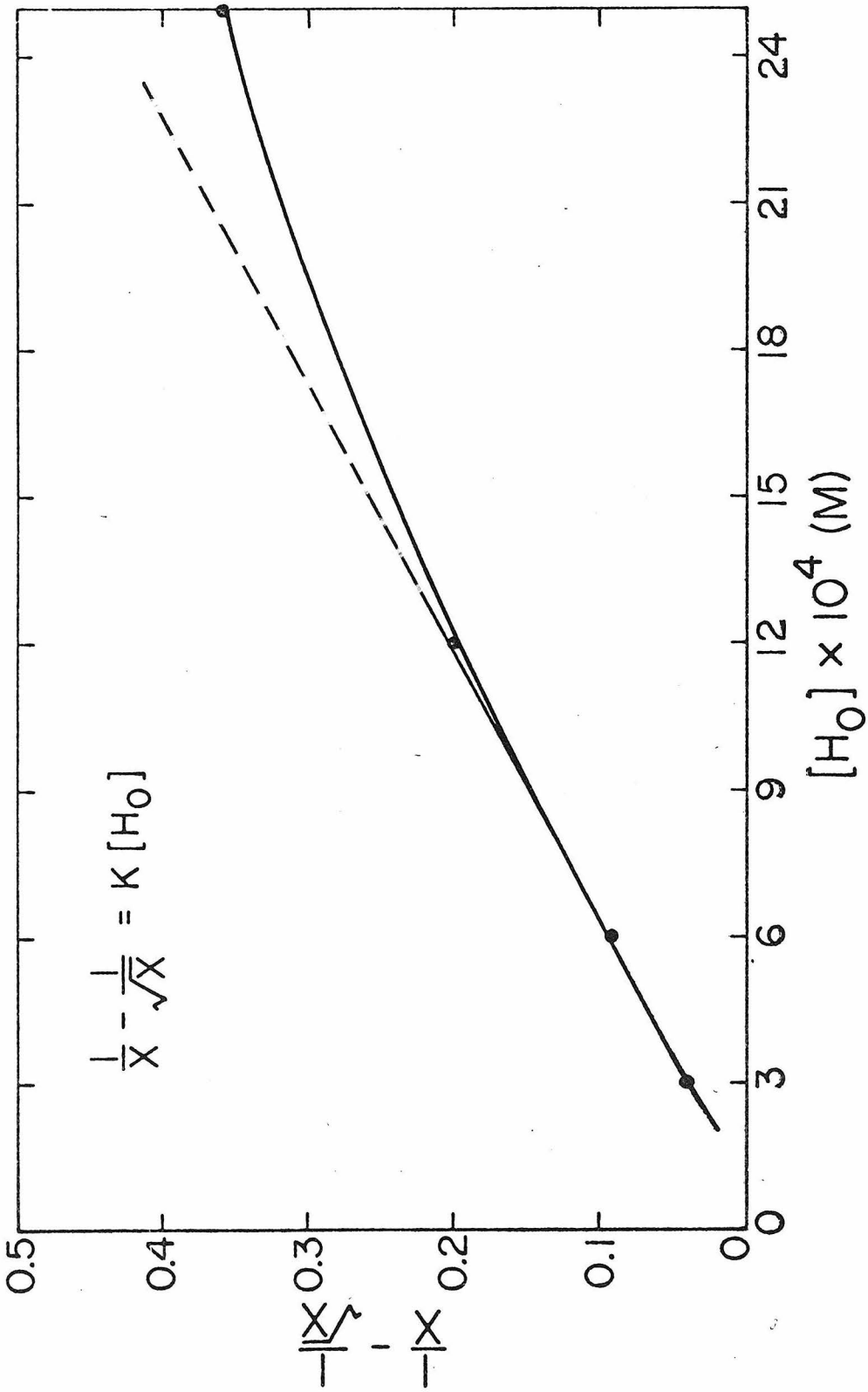

0.5
0.4
0.3
0.2
0.1
0
0
3
6
9
12
15
18
21
24
$\frac{1}{X} - \frac{1}{\sqrt{X}}$
$\frac{1}{X} - \frac{1}{\sqrt{X}} = K[H_O]$
$[H_O] \times 10^4$ (M)

Figure 15. Plot of $1/x - 1/\sqrt{x}$ vs. $[H_o] \times 10^4$ (Model 2, histone IIb1) where x is the fraction of histone IIb1 in the free state and $[H_o]$ represents the stoichiometric concentration of histone IIb1.

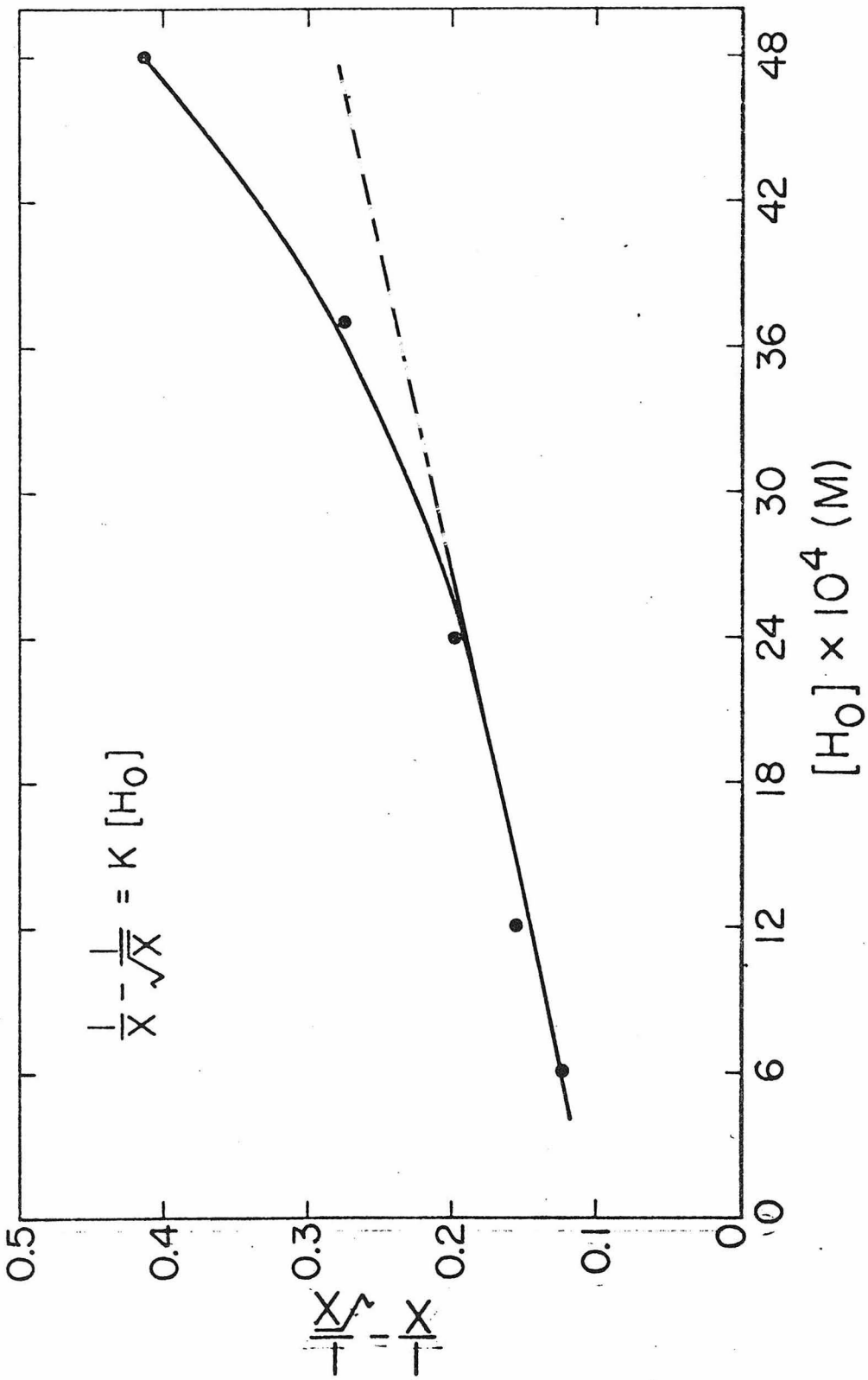

0.5
0.4
0.3
0.2
0.1
0
0
6
12
18
24
30
36
42
48
$\frac{1}{X} - \frac{1}{\sqrt{X}} = K\,[H_0]$
$\frac{X/\sqrt{X}}{1} = \frac{X}{1}$
$[H_0] \times 10^4$ (M)

Figure 16. Plot of $1/x - 1/\sqrt{x}$ vs. $[H_o] \times 10^4$ (Model 2, histone I 0.5 M NaCl), where x is the fraction histone I in the free state and $[H_o]$ is the stoichiometric concentration of histone I.

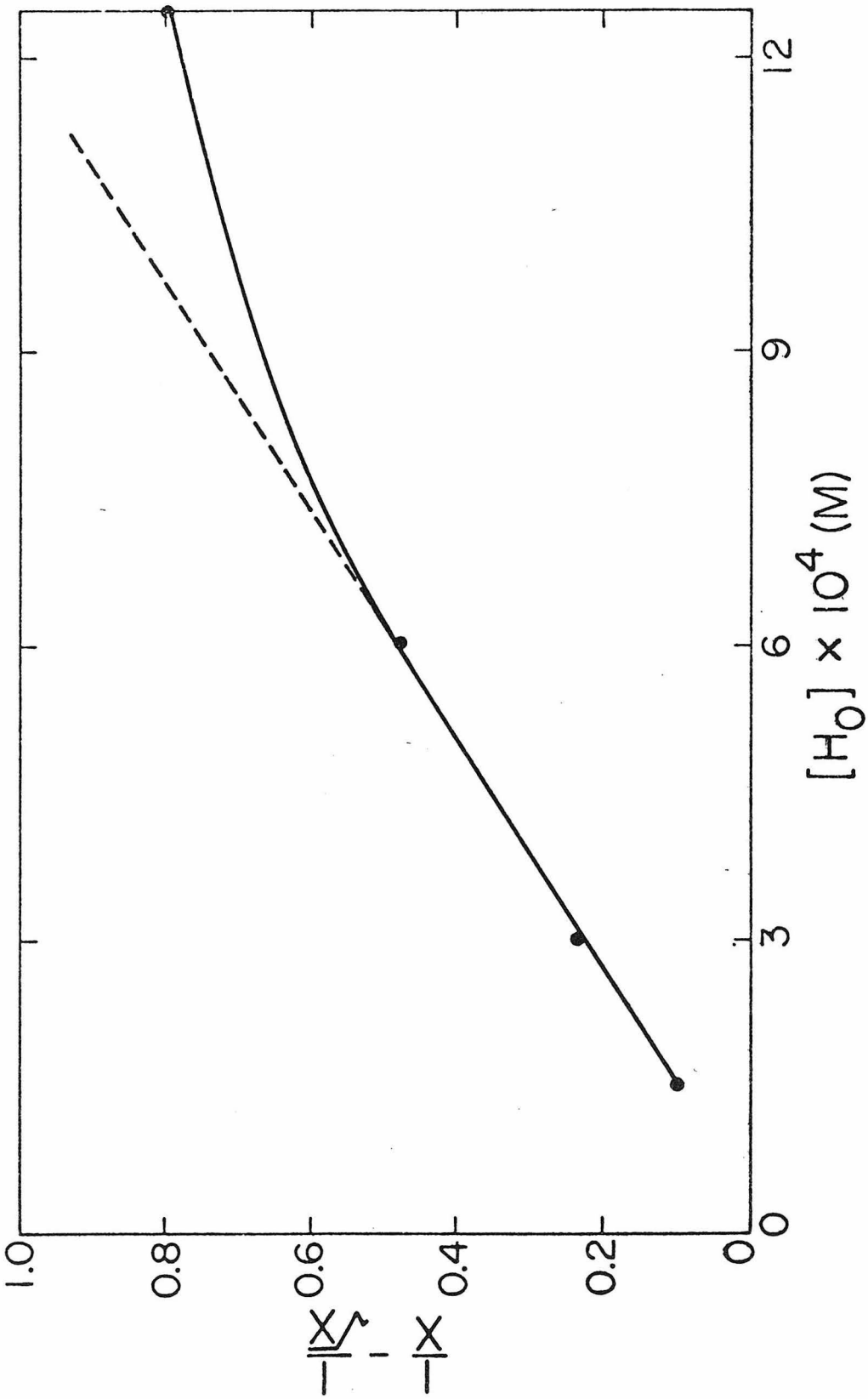
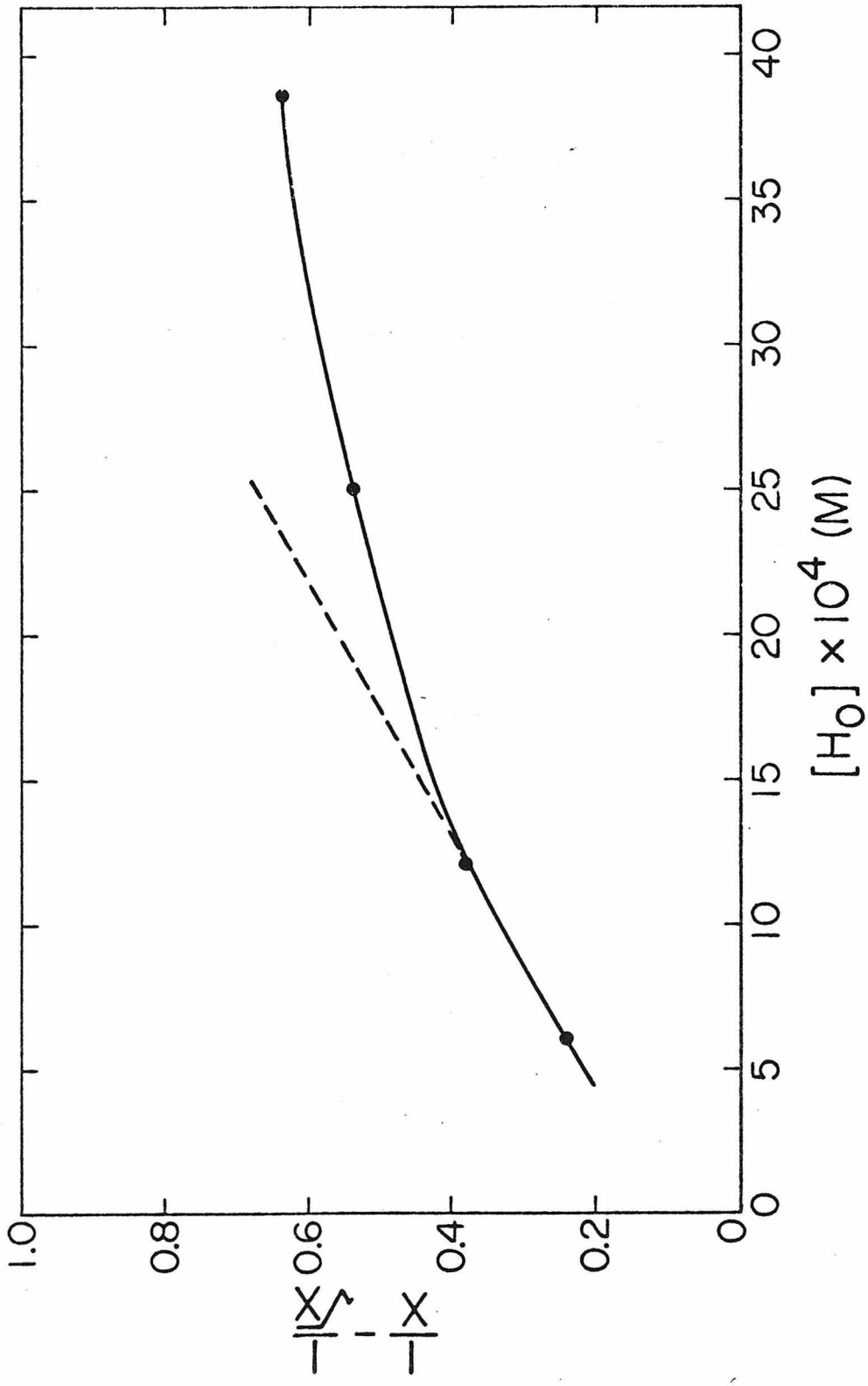

1.0
0.8
0.6
0.4
0.2
0
0
3
6
9
12
$[H_0] \times 10^4$ (M)
$\frac{X'}{1-X'} - \frac{X}{1-X}$

Figure 17. Plot of $1/x - 1/\sqrt{x}$ vs. $[H_o] \times 10^4$ (Model 2, histone IIb1 0.15 M NaCl), where x is the fraction of histone IIb1 in the free state, and $[H_o]$ is the stoichiometric concentration of histone IIb1.


1.0
0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
35
40
$\frac{\bar{X}}{1-\bar{X}} - \frac{X}{1-X}$
$[H_0] \times 10^4$ (M)

If this model is adequate to describe the aggregation of histone molecules, the plot of $1/X - 1/\sqrt{X}$ vs. $[H_o]$ should be a straight line with slope K. However, such a linear plot cannot be obtained from our experimental data either in the presence or in the absence of salt. This would indicate that the equilibrium constants for the higher aggregates are not the same as that of the dimer formation. (Fig.14 to Fig. 17).

Model 3. (polymeric model with dimer equilibrium constant K_2 and the identical equilibrium constant K for the higher aggregates)

This model states that the equilibrium constant for the dimer formation is different from those of higher aggregates and the steps for higher aggregations have the same equilibrium constants. This is best illustrated by the following equilibrium steps.

$$H_1 \overset{K_2}{\rightleftharpoons} H_2 \overset{K_3}{\rightleftharpoons} H_3 \rightleftharpoons \overset{K_n}{\rightleftharpoons} H_n$$

$$K_2 \neq K_3 = K_4 \cdots = K_n = K$$

If we assume that only the monomers contribute to the observed pmr spectra, i.e., $[H_1] \overset{\sim}{\sim} [H]_{obs}$, the total histone concentration $[H_o]$ can be obtained in terms of the monomer concentration

$$[H_o] = [H_1] + 2\,[H_2] + \cdots\cdots + n\,[H_n].$$

$$= [H_1] + [H_2]\,\{2 + 3\,[H_1]\,K. \cdots\cdots + n(K[H_1])^{n-2}\}$$

$$= [H_1] + K_2[H_1]^2 \sum_{n=2}^{\infty} n\,(K[H_1])^{n-2}$$

$$= [H_1] + K_2[H_1]^2 \cdot \frac{2 - K[H\ ]}{1 - K[H\ ]^2}$$

$$[H_o] = [H]_{obs} + K_2[H]^2_{obs}\,\frac{2 - K[H]_{obs}}{(1 - K[H]_{obs})^2}$$

$$X = [H]_{obs}/[H_o]$$

$$[H_o] = [H_o]\,X + K_2[H_o]^2\,X^2 \cdot \frac{2 - K[H_o]\,X}{(1 - K[H_o]\,X)^2}$$

$$\frac{1}{K_2[H_o]}\left(\frac{1}{X^2} - \frac{1}{X}\right) = \frac{2 - K[H_o]\,X}{(1 - K[H_o]\,X)^2}$$

$$\left(\frac{1}{X^2} - \frac{1}{X}\right) = 2\,K_2[H_o]\left[\frac{1 - \frac{K}{2}[H_o]\,X}{(I - K[H_o]\,X)^2}\right]$$

According to this model, at low histone concentration $[H_o]K \cdot X << 1$, we should get a linear plot of $(1/X^2 - 1/X)$ vs. $[H_o]$ with slope $2K_2$. However, at high concentrations when $K[H_o]\,X$ is increased, the formation of higher aggregates becomes important. Then the plot will deviate from this straight line.

Figure 18. Plots of $1/x^2 - 1/x$ vs. $[H_o] \times 10^4$ (Model 3, histone I) x is the fraction of histone I in the free state and $[H_o]$ is the stoichiometric concentration of histone I. (data extracted; spectral region I● region III + IV ▲ and region VII + VIII■)

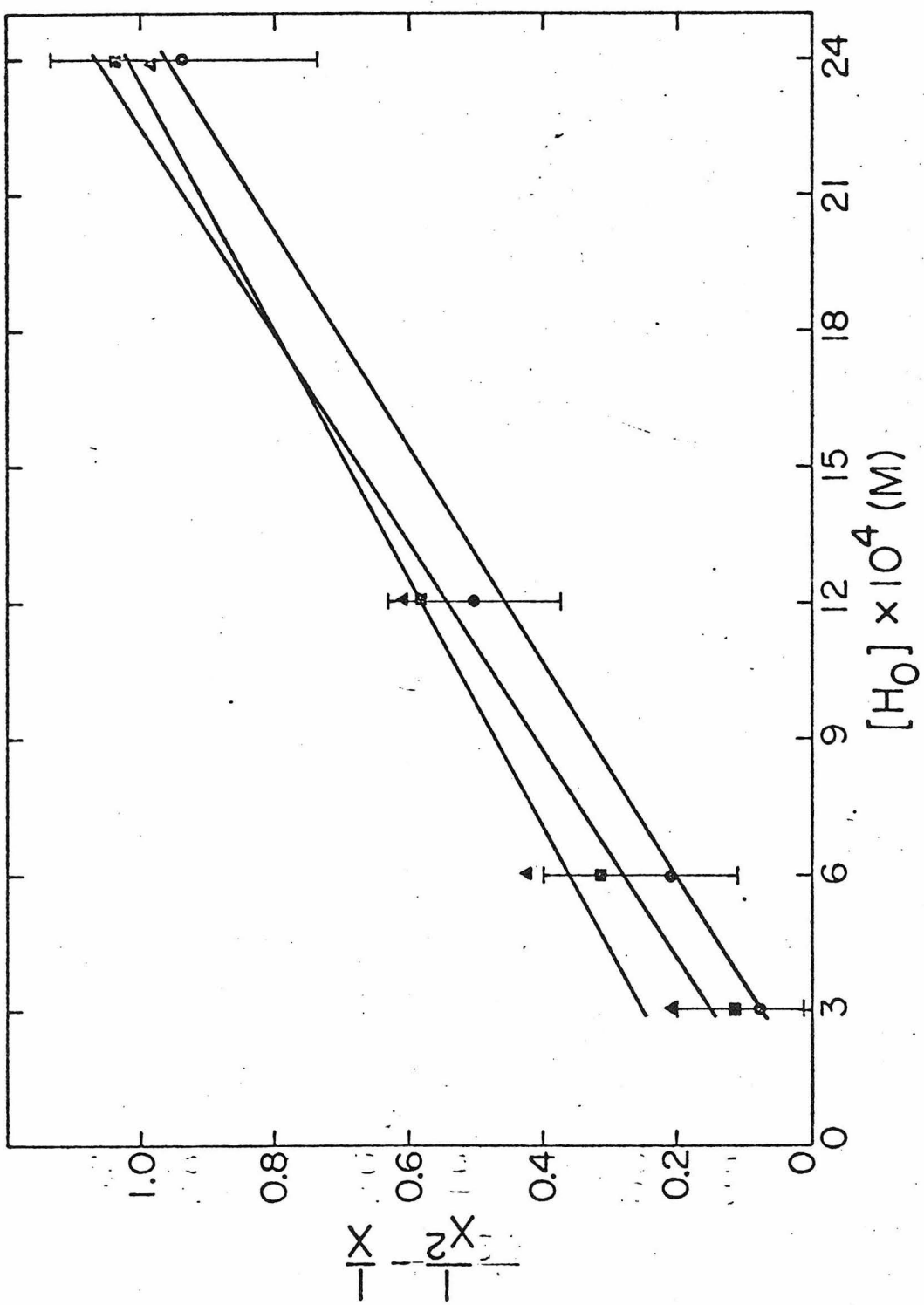

0
0.2
0.4
0.6
0.8
1.0
$\frac{X}{1-X^2}$
0
3
6
9
12
15
18
21
24
$[H_0] \times 10^4$ (M)

Figure 19. Plots of $1/x^2 - 1/x$ vs. $[H_o] \times 10^4$ (Model 3, histone IIb1) x is the fraction of histone IIb1 in the free state and $[H_o]$ is the stoichiometric concentration of histone IIb1. (data extracted; spectral region I, ● region III + IV, ▲ region VII + VIII, ■)

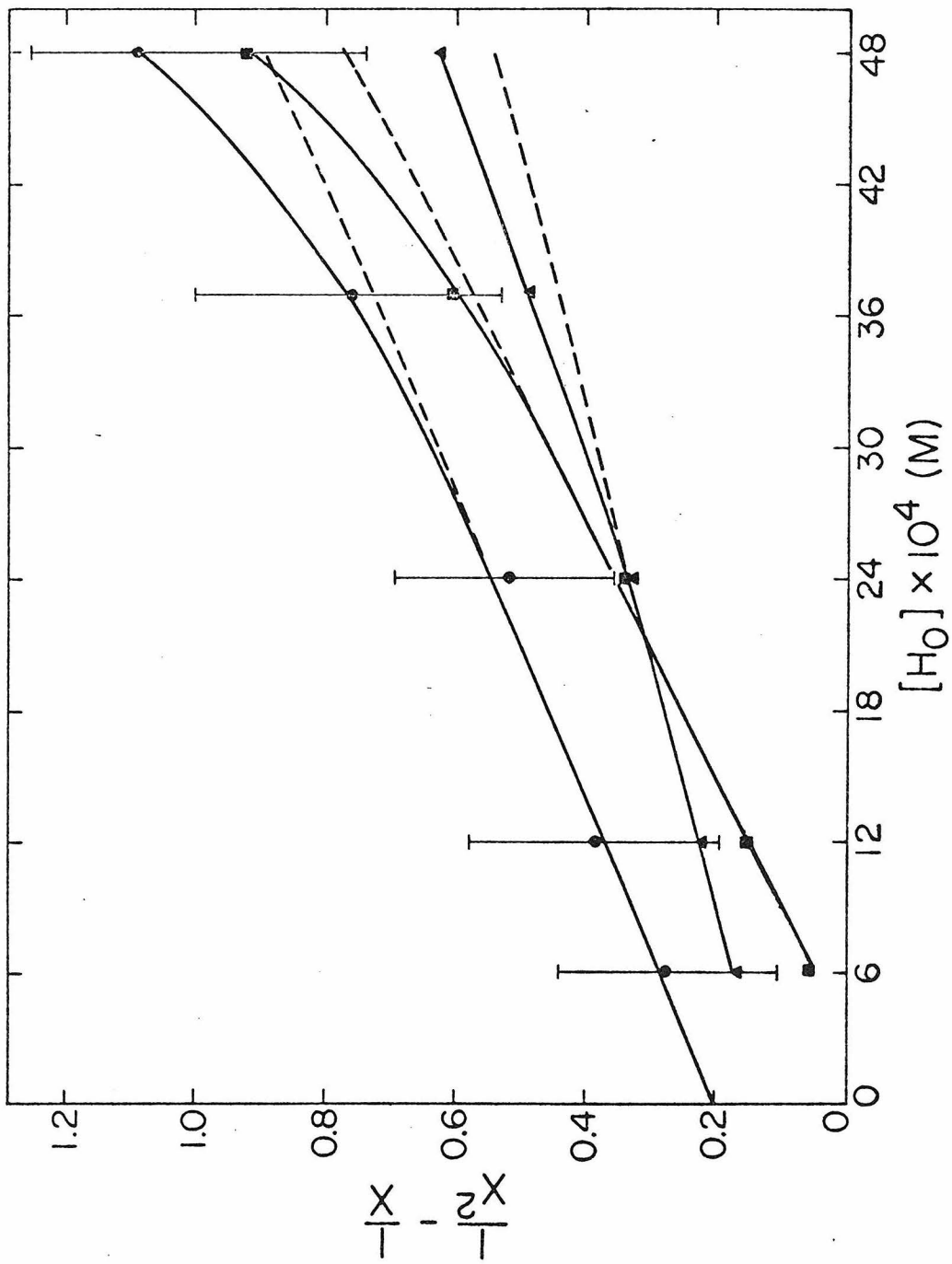

$\frac{1}{\chi^2} - \frac{1}{\chi}$
0
0.2
0.4
0.6
0.8
1.0
1.2
0
6
12
18
24
30
36
42
48
$[H_0] \times 10^4$ (M)

Figure 20. Plots of $1/x^2 - 1/x$ vs. $[H_o] \times 10^4$ (Model 3, histone I, 0.5 M NaCl), x is the fraction of histone I in the free state and $[H_o]$ is the stoichiometric concentration of histone I. (data extracted; spectral region I●, region III + IV ▲, region VII + VIII ■.)

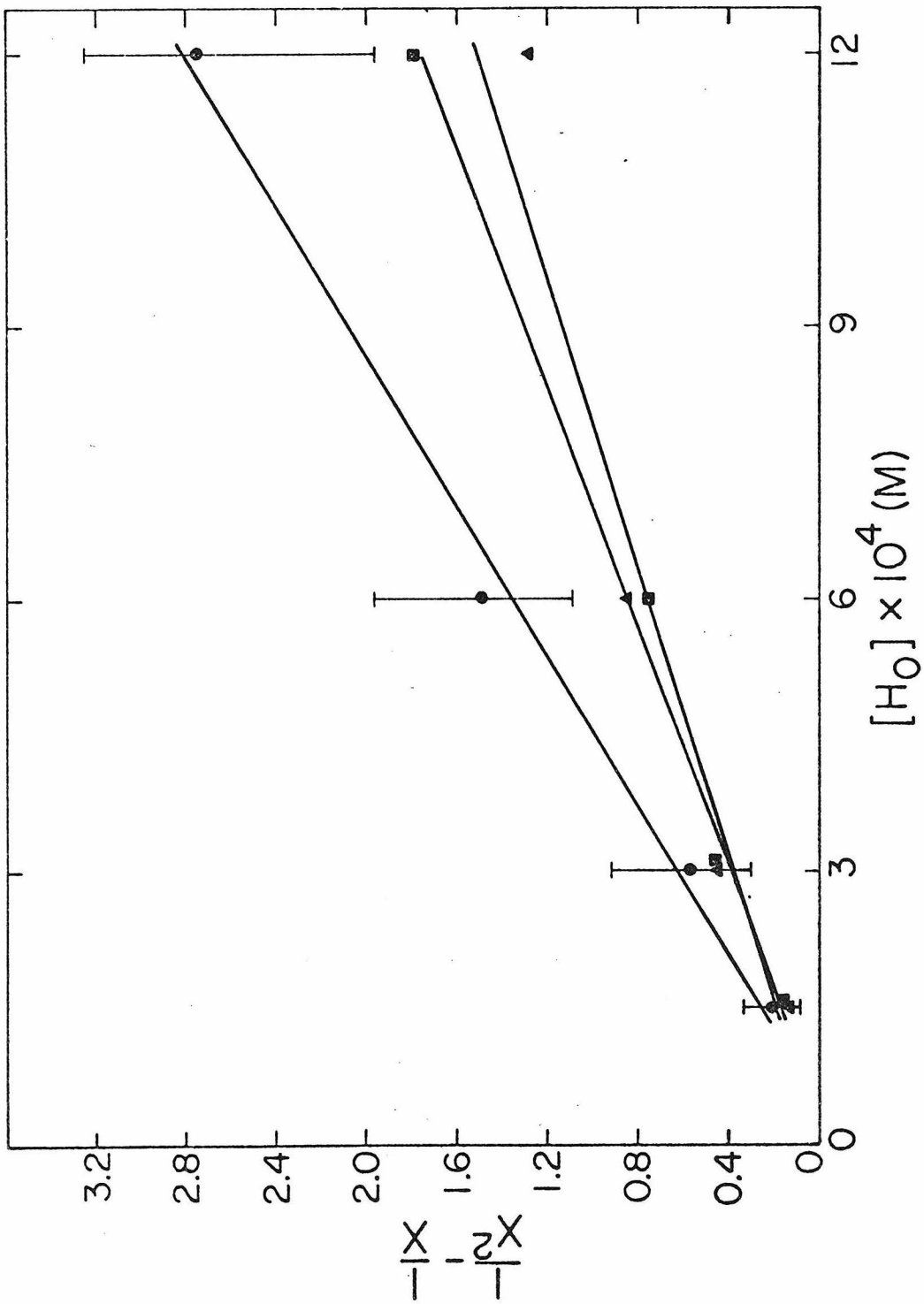
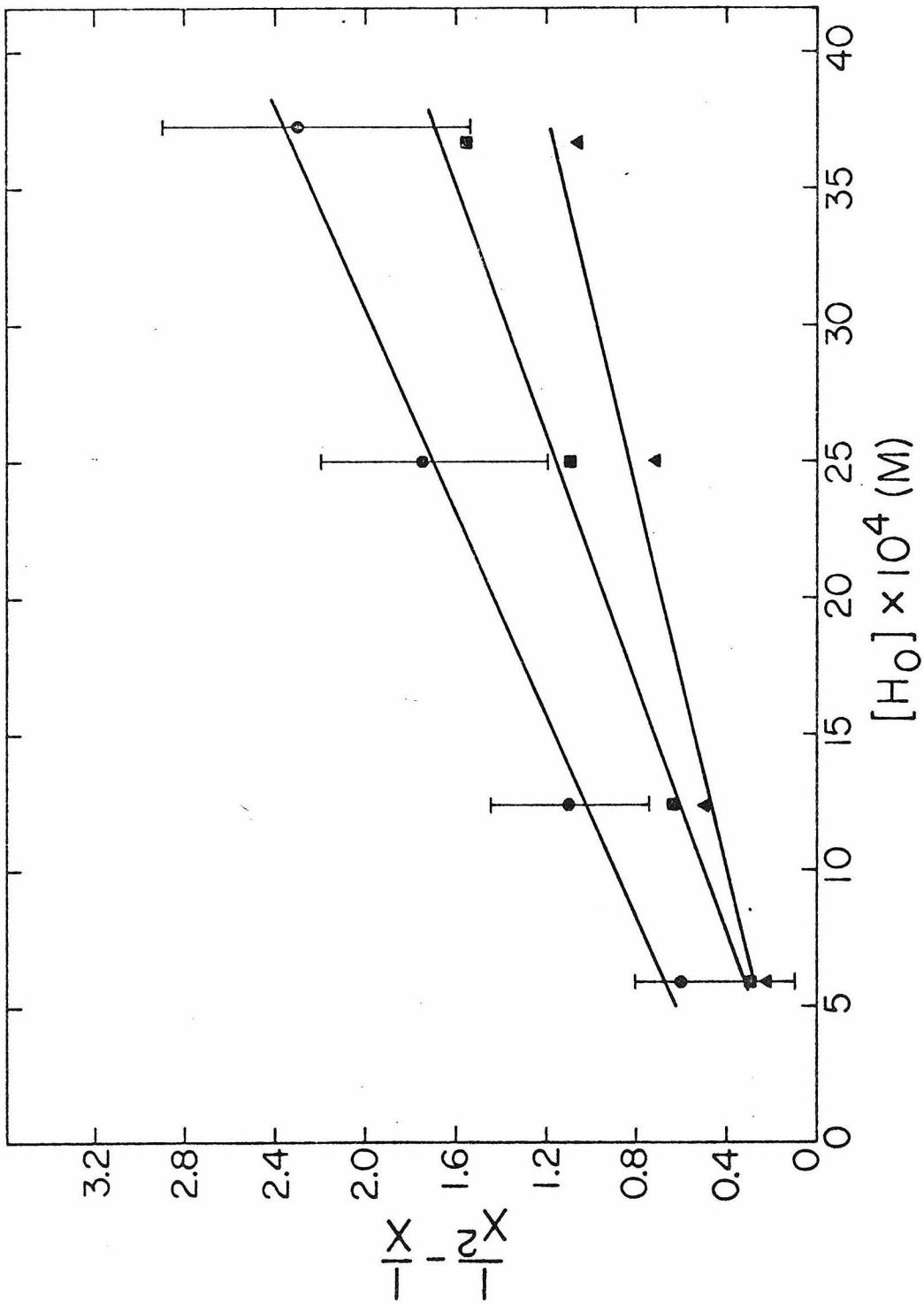

3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
0
3
6
9
12
$[H_0] \times 10^4$ (M)
$\frac{1}{\bar{X}} - \frac{2}{\bar{X}}$

Figure 21. Plots of $1/x^2 - 1/x$ vs. $[H_o] \times 10^4$ (Model 3, histone IIb1, 0.15 M NaCl), x is the fraction of histone IIb1 in the free state and $[H_o]$ is the stoichiometric concentration of histone IIb1. (data extracted; spectral region I●, region III + IV ▲, region VII + VIII ■)


$\frac{1}{X^2} - \frac{1}{X}$
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
0
5
10
15
20
25
30
35
40
$[H_0] \times 10^4$ (M)

In the case of histone I, both in the absence of salt and in the presence of 0.15 M NaCl, (Figs. 18 and 20) two linear plots were, indeed, obtained with slopes $K_2 \sim 200$ and ~ 900, respectively. This would indicate that only the dimer is formed within the concentration range (3 mg/ml to 50 mg/ml). In this situation $K \tilde{\sim} 0$ and the equation for the dimer formation $1/X^2 - 1/X = 2\ K_2[H_o]$ can, therefore, be applied. However, for histone IIb1 in the salt free solution, the plot does not show a straight line at high concentrations. This indicates that the term $K[H_o]X$ becomes important at high concentrations and the formation of the higher aggregates becomes important. The best fit between the theoretical equation and the experimental curve shows that $K_2 \tilde{\sim} 100$ $K \tilde{\sim} 20$ for this equation. While in the presence of 0.15 M NaCl the dimer formation still predominates up to the concentration 25 mg/ml with $K \sim 250$. (Figs. 19 and 21).

(3). Effect of Salt on Histone Aggregation

Based on the results of the model studies on histone I and histone IIb1, we have concluded that the histone dimer is the favorable aggregated species either in the presence of or in the absence of salt in neutral aqueous solution. The condensations of the counter ions to the charged amino acid residues of these proteins are known to be the major driving force for the histone aggregations. In order to understand this mechanism, it is desirable to obtain quantitative thermodynamic relations between the equilibrium of the histone aggregations and the activity of the associated solution components such as the salt and the solvent.

(4). Generalized Theory[14,15]

The association of the solvent and salt molecules to a macromolecule in the solution can generally be expressed by the following equation.

$$M + iX + jW \rightleftarrows M X_i W_j \quad ,$$

where M represents the macromolecule, X is the ligand or counter ion of a macromolecule, W is the solvent molecule-water, i and j represent the number of ligand and water molecules associated to a macromolecule, respectively, resulting in the formation of a particular species MX_iW_j. According to Wyman, if the associations involve many steps with the formation of numerous species MX_iW_j and with equilibrium constant K_{ij}, then the average number of molecules X and W bound to each macromolecule $\bar{\nu}_x$ and $\bar{\nu}_w$ can be expressed by:

$$\bar{\nu}_x = \left| \frac{\partial \ln \Sigma\Sigma K_{ij} a_w^i a_x^j}{\partial \ln a_w} \right|_{a_x} \tag{5}$$

$$\bar{\nu}_w = \left| \frac{\partial \ln \Sigma\Sigma K_{ij} a_w^i a_x^j}{\partial \ln a_x} \right|_{a_w} \tag{6}$$

where a_w and a_x represent the activity of water and ligand, respectively. If a_x is chosen as the independent variable, the experimentally measurable quantity, at constant temperature and pressure, we have the following differential relations:

$$\frac{d\ \ell n\ \Sigma\Sigma K_{ij} a_w^i a_x^j}{d\ \ell n\ a_x} = \frac{\partial\ \ell n\ \Sigma\Sigma K_{ij} a_w^i a_x^j}{\partial\ \ell n\ a_x} + \frac{\partial\ \ell n\ \Sigma\Sigma K_{ij} a_w^i a_x^j}{\partial\ \ell n\ a_w} \frac{d\ \ell n\ a_w}{d\ \ell n\ a_x} \qquad (7)$$

The relation between $\ell n\ a_w$ and $\ell n\ a_x$ is given by the Gibbs-Duhem equation which for a three component system at constant temperature and pressure is

$$\ell n\ a_w = -\frac{n_x}{n_w} d\ \ell n\ a_x - \frac{n_M}{n_w} d\ \ell n\ a_M \quad , \qquad (8)$$

where n_M, n_x, and n_w represent the molarity of macromolecules, ligand, and water in solution, respectively. Because the binding relations apply to infinite dilution of macromolecules M (i.e., $n_M = 0$), we can neglect the last term of eq. (8), so that

$$d\ \ell n\ a_w = -\frac{n_x}{n_w} d\ \ell n\ a_x$$

Combining eqs. (7) to (8), we obtain the final result.

$$\frac{d\ \ell n\ \Sigma\Sigma\ K_{ij}\ a_w^i\ a_x^j}{d\ \ell n\ a_x} = \bar{\nu}_x - \frac{n_x}{n_w} \bar{\nu}_w \qquad (9)$$

(5). The Effect of a Ligand on Macromolecular Equilibrium

Consider three macromolecules A, B, and C which react one another by the process

$$A + B \rightleftarrows C$$

Let K be the equilibrium constant for this reaction, in concentration units. If we assume that each of macromolecules can react with a ligand X and with water, we can apply the above equations to each macromolecule separately with equilibrium constant K_{ij} A, K_{ij}B, K_{ij}C.

Then the total concentration of all the A molecules $[A]_o$ can be expressed in terms of the concentration of A molecules without any bound ligand or bound water, $[A]_f$:

$$[A]_o = [A]_f \; \Sigma\Sigma \; K_{ij}A \; a_w^i \; a_x^j \tag{10}$$

Similar relations apply for B and C.

The equilibrium constant for the reactions between macromolecules becomes

$$K = \frac{[C]_f}{[A]_f \; [B]_f} \; \frac{\Sigma\Sigma \; K_{ijC} \; a_w^i \; a_x^j}{\Sigma\Sigma \; K_{ijA} \; a_w^i \; a_x^j \; \Sigma\Sigma \; K_{ijB} \; a_w^i \; a_x^j} \; . \tag{11}$$

$\frac{[C]_f}{[A]_f \; [B]_f}$ represents the equilibrium constant for the reaction between completely free macromolecules. Now the effect of the presence of the ligand X on the equilibrium constant K is given by

$$\frac{d\ \ell n\ K}{d\ \ell n\ a_x} = \frac{d\ \ell n\ \Sigma\Sigma_C}{d\ \ell n\ a_x} - \frac{d\ \ell n\ \Sigma\Sigma_A}{d\ \ell n\ a_x} - \frac{d\ \ell n\ \Sigma\Sigma_B}{d\ \ell n\ a_x} \qquad (12)$$

where $\Sigma\Sigma_A$, $\Sigma\Sigma_B$, and $\Sigma\Sigma_C$ denote the appropriate sums of eq. (10). Substituting eqs. (5) and (6) into eq. (12) gives

$$\frac{d\ \ell n\ K}{d\ \ell n\ a_x} = \bar{\nu}_{x,C} - \bar{\nu}_{x,A} - \bar{\nu}_{x,B} - \frac{n_x}{n_w}(\bar{\nu}_{w,C} - \bar{\nu}_{wA} - \bar{\nu}_{wB})$$

$$= \Delta\bar{\nu}_x - \frac{n_x}{n_w}\Delta\bar{\nu}_w$$

where $\bar{\nu}_{x,A}$, $\bar{\nu}_{x,B}$, and $\bar{\nu}_{x,C}$ represent the average number of ligands associated with macromolecules A, B, and C, respectively. $\bar{\nu}_{w,A}$, $\bar{\nu}_{w,B}$, and $\bar{\nu}_{w,C}$ represent the average number of hydration of the respective species. $\Delta\bar{\nu}_x$, and $\Delta\bar{\nu}_w$ represent the change in number of associated ligand or hydration due to the association of macromolecules. At low ligand concentrations $n_x << n_w$ we have

$$\frac{d\ \ell n\ K}{d\ \ell n\ a_x} \approx \Delta\bar{\nu}_x \qquad (13)$$

Equation (12) can be applied to our system, if we assume that salt dependent histone aggregations are due to equilibrium between histone dimers and monomers with equilibrium constant K.

$$2\ H \rightleftarrows H_2 \quad ,$$

where H and H_2represent histone monomer and dimer, respectively.

The equilibrium constant K at different salt activities can be obtained from the salt dependence of the spectral area loss of histone I and histone IIb1. (Fig. 22). In this study, we pick up Region I to estimate the equilibrium constant K at different salt concentration.

At low salt activity (or counter ion activity), a straight line can be obtained from the plot of ℓn K vs. ℓn a_x with slope $\Delta\bar{\nu}_x$ where a_x is the activity of salt or counter ions. In Fig. 22, however, this plot is modified by expressing ℓn K in terms of the free energy change (ΔG) of histone dimerization. From $\Delta G = -RT$ ℓn K, we have

$$\frac{d[-\Delta G]}{d \log a_x} 2.303\ RT \frac{d \log K}{d \log a_x} = 2.303\ RT\ \Delta\bar{\nu}_x$$

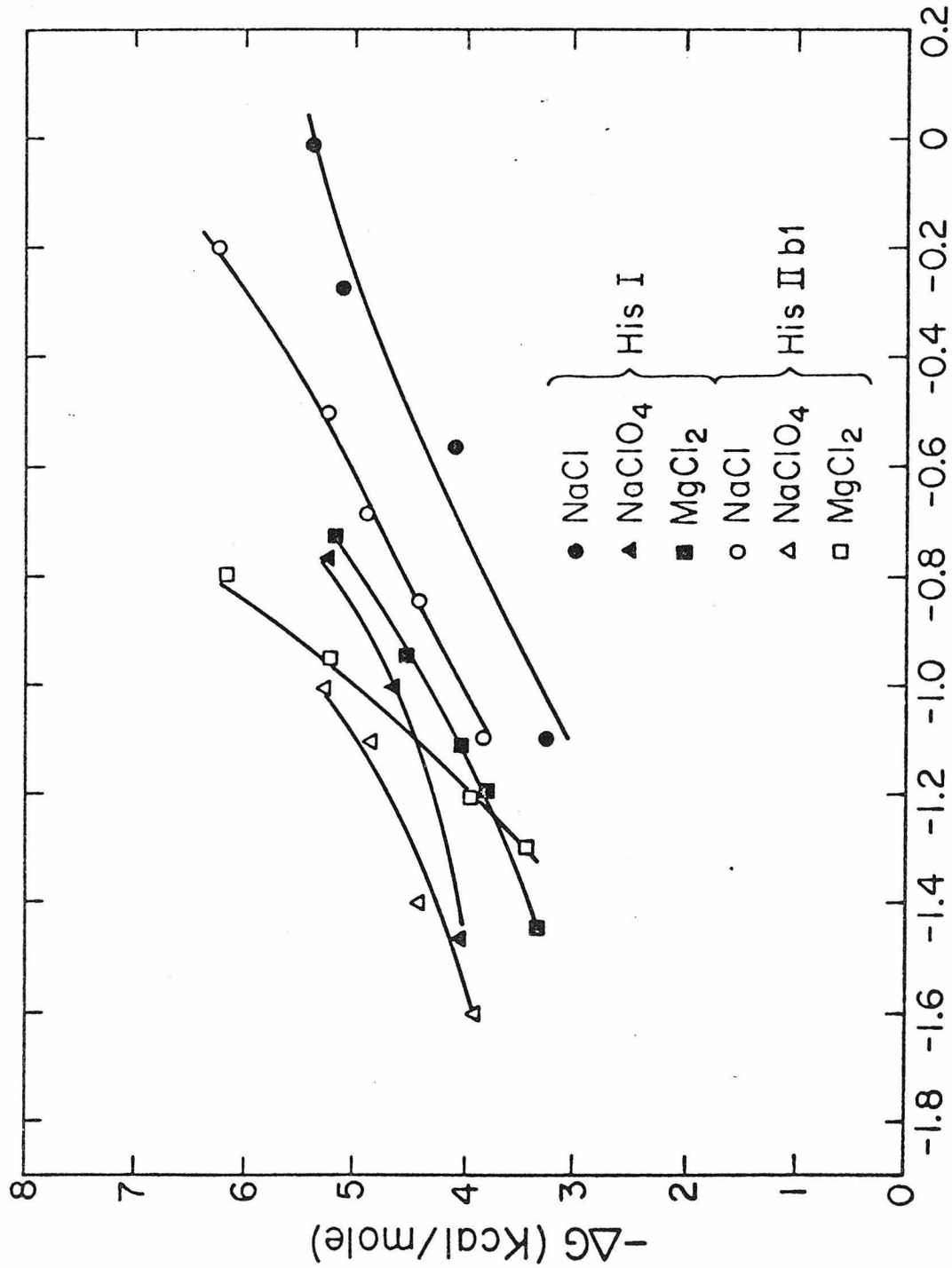
If we plot $-\Delta G$ vs. log a_x we can also get a straight line with slope 2.303 RT $\Delta\bar{\nu}_x$ from which we can estimate the change in the number of counter ion association due to histone aggregation $\Delta\bar{\nu}_x$. $\Delta\bar{\nu}_x$'s estimated from different salt (NaCl, $MgCl_2$ and $NaClO_4$) are given in Table V.

At high salt concentrations, the contribution from the changes of water hydration due to histone aggregation cannot be neglected. Hence, from eq. (12), we have

$$\frac{d(-\Delta G)}{d \log a_x} = 2.303\ RT \frac{d \log K}{d \log a_x} = 2.303\ RT(\Delta\bar{\nu}_x - \frac{n_x}{n_w} \Delta\bar{\nu}_w)$$

where n_x and n_w denote the molarity of salt and water molecules, respectively. $\Delta\bar{\nu}_w$ is the change of water hydration due to dimer formation.

Figure 22. Plot of - ΔG vs. log $\mathfrak{a}$, where ΔG is the free energy change for the dimer formation of histones and $\mathfrak{a}$ is the activity of the salt.


-ΔG (Kcal/mole)
LOG a
NaCl
NaClO4
MgCl2
NaCl
NaClO4
MgCl2
His I
His II b1

$\Delta\bar{\nu}_w$ can be estimated from the slope of the plot of log K vs. log a_x at highest salt concentrations. $\Delta\bar{\nu}_w$ for different salts are given in Table V.

From this study, we may conclude that the counter ion binding is increased in the histone dimers ($\Delta\nu_x > 0$) and the equilibrium constant for the dimer formation is increased with increasing salt concentration.

Histone I and histone IIb1 are known to have high content of basic amino acid residues compared to the acidic amino acid residues. This would result in the asymmetric charge distribution of these protein molecules. The presence of salt will certainly increase the counter ion binding to the proteins and the deshielding of long range electrostatic interaction between protein molecules is known to occur. As a result, the aggregation between the protein molecules due to the hydrophobic contact between the hydrophobic side chain of the amino acid residues is enhanced to the point of increase in spectral area loss with increasing salt concentration. The changes in hydration number of protein molecules due to molecular aggregation is also significant. This could be due to the conformational changes of the protein upon the intermolecular aggregation.

In order to understand more about the physical picture of this protein aggregation from thermodynamic point of view, we study the spectral area loss of histone I as well as histone IIb1 as a function of temperature both in the presence and in the absence of salt.

TABLE V

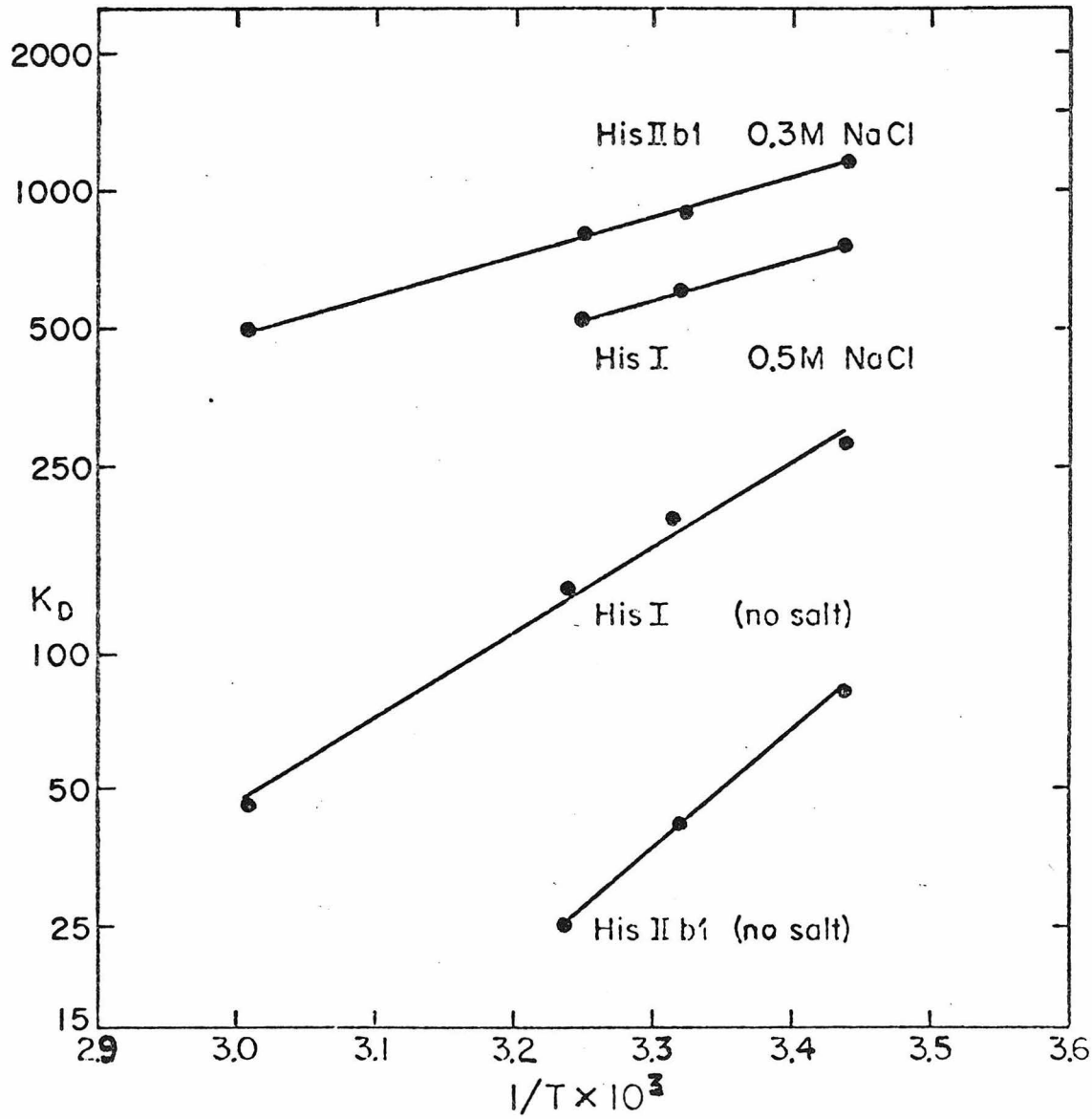
		His I	His IIb1
NaCl	$\Delta\bar{\nu}_x$	1.9	1.8
	$\Delta\bar{\nu}_w$	25	-30
$NaClO_4$	$\Delta\bar{\nu}_x$	0.9	1.5
	$\Delta\bar{\nu}_w$	-60	-100
$MgCl_2$	$\Delta\bar{\nu}_x$	1.1	3.6
	$\Delta\bar{\nu}_w$	87	-30

The logarithm of equilibrium constant K estimated from the spectral area changes with temperature is plotted against the inverse of temperature. The approximate enthalpy and entropy changes due to the dimer formation can then be estimated from this plot (ΔH and ΔS).

It is interesting to know that in the presence of salt, the enthalpy changes (ΔS) due to dimer formation become more positive compared to those in the absence of salt. For example, histone I and histone IIb1 in salt free solution exhibit $\Delta S = -12$ e.u. and -29 e.u., respectively, for dimer formation. While in the presence of 0.5 M NaCl histone I exhibits $\Delta S = +0.35$ e.u. and histone IIb1 in 0.3 M NaCl solution exhibits $\Delta S = +0.24$ e.u. The dramatic effect of salt on the hydrophobic protein aggregation is clearly indicated from these thermodynamic quantities as expected. (Fig. 23).

The observed spectral area of histone I and histone IIb1 is also pD dependent especially at extreme pD's. This phenomenon can also be interpreted qualitatively based on the mechanism of protein aggregation. At pD between 9 to 12, the deprotonation of the ammonium group of lysine residues as well as guanidine group of arginine residues will decrease the electrostatic interaction between the protein molecules and thus promote the aggregation by hydrophobic interaction. According to Tanford's equation, the effect of the activity of the acid, base, as well as solvent on the apparent equilibrium constant K for protein aggregation can be expressed in the following (assuming in salt free solution)

Figure 23. Semi-log plot of K_D vs. $1/T \times 10^3$, where K_D is the equilibrium constant for the dimer formation of histones.


2000
1000
500
250
100
50
25
15
K_D
His II b1 0.3M NaCl
His I 0.5M NaCl
His I (no salt)
His II b1 (no salt)
2.9
3.0
3.1
3.2
3.3
3.4
3.5
3.6
$1/T \times 10^3$

$$d \ln K = \left[\frac{\partial \ln K}{\partial \ln a_A} \right]_{a_W, a_B} d \ln a_A + \left[\frac{\partial \ln K}{\partial \ln a_B} \right]_{a_W, a_A} d \ln a_B$$

$$+ \left[\frac{\partial \ln K}{\partial \ln a_W} \right]_{a_A, a_B} d \ln a_W$$

where a_A, a_B, and a_W are the respective activity of acid, base, and water.

At pD 9.0 to 12, we may assume that the quantities $\frac{d \ln a_W}{d \ln a_B}$ and $\frac{d \ln a_A}{d \ln a_B}$ are negligibly small, so that

$$\therefore \frac{d \ln K}{d \ln a_B} = \left[\frac{\partial \ln K}{\partial \ln a_B} \right]_{a_A, a_W}$$

According to Wyman's relation, we have

$$\frac{d \ln K}{d \ln a_B} = + \frac{d \log K}{d pD} = \Delta \bar{\nu}_B = \bar{\nu}_B^{(n)} - n \cdot \bar{\nu}_B^{(1)}$$

where $\Delta \bar{\nu}_B$ is the difference in base binding between the equilibrium species, $\bar{\nu}_B^{(n)}$, $\bar{\nu}_B^{(1)}$ are the base binding of aggregated species and monomer, respectively. n represents the average number of aggregations.

If the observed area loss at high pD's properly indicates the molecular aggregation, then the changes in equilibrium constant K

due to pD changes would be directly related to the difference in base association with protein among the equilibrium species. The highest degree of molecular aggregation at pI (Isoelectric point) of these proteins is indicated by the maximum area loss at the corresponding pD. At pD higher than the isoelectric point, the ionization of αNH group of the amino acid residues of the proteins would increase the charge repulsion between the protein molecules and thus decrease the equilibrium constant for the aggregation as indicated by the rapid increase in the spectral area at pD > pI.

Based on the same argument, the increase in protein aggregation indicated by the spectral area loss at low pD's is clearly related to the changes in acid association $\Delta\bar{\nu}_A$ between the equilibrium species by $\frac{d \log K}{d\ pD} = -\Delta\bar{\nu}_A$, where K is the equilibrium constant for the protein aggregation.

5. CONCLUSION

We have presented our pmr study of nucleoproteins - histone I and histone IIb1 at various protein concentrations, salt concentrations as well as different pD's. The observed progressive spectral area loss with increasing protein or salt concentrations was interpreted as the result of increasing protein aggregations. The quantitative treatments between the spectral area loss and various types of protein aggregations were discussed with three different models. It was concluded that the

dimer is the most favorable species for these two basic proteins either in the presence or in the absence of salts within the concentration range 3 mg/ml to 50 mg/ml.

The enhancement of the protein aggregation in the presence of salt is primarily due to the changes in the entropy difference (ΔS) between the aggregated and free species. From the temperature study, it was deduced that in the presence of salt ΔS becomes more positive than that in the absence of salt, as a result of the increase in the hydrophobic interaction between the protein molecules.

REFERENCES

1. R. C. C. Huang and J. Bonner, Proc. Nat. Acad. Sci. U.S. 48, 1216 (1962).
2. G. C. Barr and J. A. V. Butler, Nature 199, 1170(1963).
3. V. G. Allfrey, V. D. Littau, and A. E. Mirsky, Proc. Nat. Acad. Sci. U.S. 49,414(1963).
4. D. Y. H. Tuan and J. Bonner, J. Mol. Biol., 45,59(1969).
5. P. A. Edward and K. V. Shooter, Biochem. J., 114,227(1969).
6. E. M.Bradbury, C. Crane-Robinson, H. Goldman, H. W. E. Rattle, and R. M. Stephens, J. Mol. Biol., 29,507(1967).
7. M. Boublik, E. M. Bradbury, and C. Crane-Robinson, Eur. J. Biochem. 14,486(1970).
8. M. Boublik, E. M. Bradbury, C. Crane-Robinson, and E. W. Johns, Eur, J. Biochem. 17,151(1970).
9. E. Peakery, H. J. Li, S. I. Chan, C. J. Hsu, and J. E. Wagner, Biochemistry, in press (1972).
10. E. W. Johns, Biochem, J., 92,55(1964).
11. R. Lumry, E. L. Smith, and R. R. Glantz, J. Am. Chem. Soc., 73,4335 (1951).
12. C. C. McDonald and W. D. Phillips, J. Am. Chem. Soc., 91,1513(1969).
13. A. Abragam, The Principles of Nuclear Magnetism, Oxford University Press, 1962.
14. C. Tanford, J. Mol. Biol., 39,539(1969).
15. J. Wyman, J. Mol. Biol., 11,631(1965).

II. PROTON MAGNETIC RESONANCE STUDIES OF THE DNA-HISTONE I COMPLEX

1. INTRODUCTION

Histones, when complexed with DNA, are known to repress the template or gene activity of DNA by inhibiting the process of RNA transcription[1-3]. Many attempts have been undertaken to characterize the structures and properties of these nucleoproteins and their complexes with DNA including uv, CD, ORD[4-11], and pmr studies[12-14]. Despite these efforts, our understanding of the detailed molecular structures of the DNA-histone I associated complexes is still pretty fragmentary.

This paper describes an attempt to study a reconstituted nucleo-histone by high resolution proton magnetic resonance spectroscopy (pmr). For this study, we have chosen the DNA-histone I complex. Among the various DNA-histone complexes, the DNA-histone I complex is known to have more favorable solubility characteristics in aqueous solution and is therefore more amenable to the pmr method. Also, it has been suggested that this complex has a rather loose structure and is easily dissociated in the presence of salt[1, 15]. The melting temperature of the complexed DNA is also known to be lower than those of similar complex[12] involving other histones.

2. EXPERIMENTAL

Calf thymus histone I was kindly provided by Dr. Bonner of the Division of Biology at Caltech. Calf thymus DNA was purchased from Calbiochem as grade A reagent. In order to eliminate possible protein contamination, the DNA was first purified by three successive phenol extractions. It was then precipitated from 95% ethanol and redissolved in 0.1 M NaCl solution. Finally it was dialyzed against a solution containing 0.01 M NaCl and 0.001 M tris buffer at pH 8.0.[7]

Reconstituted DNA-histone I complex was prepared by continuous salt gradient dialysis method adapted and modified from the stepwise salt gradient dialysis procedure of Huang et al.[6] DNA and histone I at various input ratios of concentrations were mixed in a solution containing 2 M NaCl and 0.01 M tris buffer at pH 8.0. NaCl was continuously dialyzed out to the concentration of 0.05 M. The mixture was then extensively dialyzed against 0.001 M tris, at pD 8.0 in D_2O. 99.7% D_2O was supplied by Columbia Organic Chemicals Co., Inc.

PMR spectra were recorded on a Varian HR-220 nmr spectrometer equipped with a Varian C-1024 time-averaging computer. A capillary containing 2,2-diphenyl-1-picrylhydrazene (DPPH), tetramethyl silane (TMS) and $DCCl_3$ served as an area as well as a chemical shift reference for all the spectra. The chemical shift of the DPPH-broadened TMS external reference was measured relative to internal DSS in a separate experiment. This frequency difference was then used to refer all chemical shifts reported in this work to "internal" DSS.

Because of the uncertainty in the base-line measurement there is a $\sim$ 10% error in the area evaluations. This error will be larger for small, poorly resolved peaks. All area measurements were made by weighing cut-outs of duplicate spectra. The capillary TMS concentration was obtained by comparing the area of the methyl resonance of TMS to the methyl resonance area of a $(CH_3)_4NCl$ solution of known concentration.

The ambient probe temperature was 17 ± 1°C in the HR-220 spectrometer. In the variable temperature studies, the sample temperature was controlled to ± 1°C by means of a variable temperature controller and was measured using the methanol and ethylene glycol standards together with calibration curves supplied by Varian.

3. RESULTS

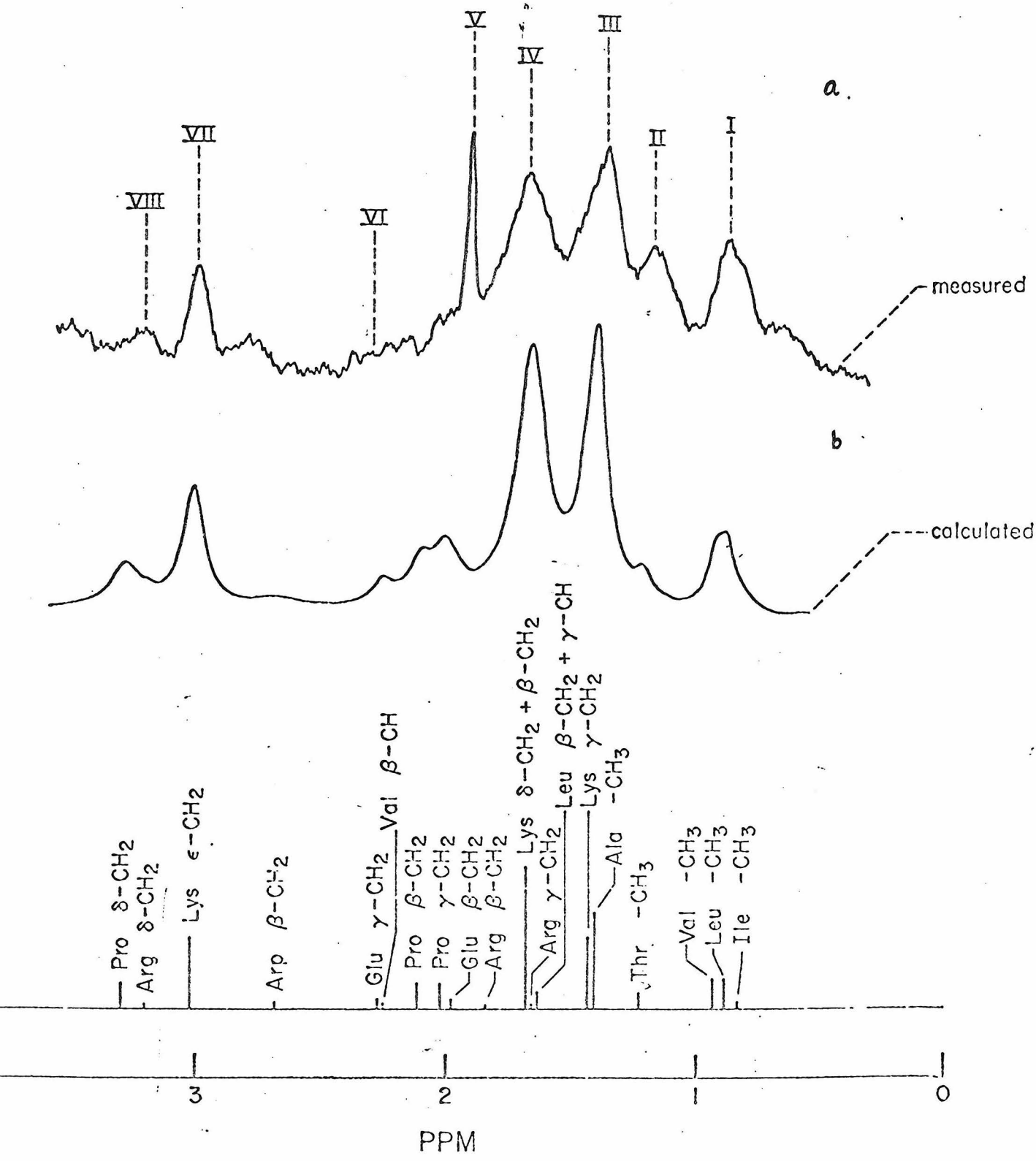
(1) PMR Spectrum of Histone I in the Free State

The 220 MHz pmr spectrum of histone I at concentration 3 mg/ml in neutral salt free aqueous solution is shown in Fig. 1(a). This spectrum reveals well defined spectral regions which may be assigned to the various amino acid residues of a random coil polypeptide. The comparison between the experimentally observed and the theoretical spectra calculated assuming a random coil conformation and known amino acid distributions is presented in Fig. 1(b).[17] The agreement between the calculated and the observed spectra supports the idea that histone I exists primarily as a random coil polypeptide under these conditions in accordance with the CD and ORD studies.

The intensity or spectral area of the histone I spectrum was found to be concentration dependent. At a protein concentration of 3 mg/ml, essentially all (∿ 95%) of the expected intensity was observed, within the experimental error. However, when the protein concentration was increased to 50 mg/ml, only 65% of the expected spectral intensity was observed. This concentration dependent area loss has been shown to arise from the intermolecular aggregations of histone I. We have studied this intermolecular association in detail. These results will be presented elsewhere.[18]

For convenience of the discussion, we have divided the high field spectrum of histone I into several spectral regions. The resonances appearing in region I + II about 1 ppm downfield from DSS, arise from the aliphatic methyl group of the aliphatic amino acid residues, Ile, Leu, Val and Thr. The resonances from methyl side chain of alanine residue as well as β, γ, δ-CH_2 group of lysine residues and γ-CH_2 group of arginine residues are accounted for by the resonances in region III and IV. Region V is due to the β-CH_2 group of glutamic acid residues and the sharp peak can be identified as α-CH_3 group of the acetic acid contaminant. Region VI includes the resonances from β, γ-CH_2 group of proline, β-CH_2 group of aspartic acid and γ-CH_2 group of glutamic acid residues. Region VII arises from the ε-CH_2 group of the lysine residue and region VIII from δ-CH_2 group of arginine and proline residues. These spectral assignments are summarized in Fig. 1(b).

Figure 1. 220 MHz histone I spectrum in the high field region at 17°C and at concentration of 3 mg/ml.[a] (The [b], computer simulated spectrum is given at the bottom).


a.
b
measured
calculated
VIII
VII
VI
V
IV
III
II
I
Pro δ-CH2
Arg δ-CH2
Lys ε-CH2
Arp β-CH2
Glu γ-CH2
Val β-CH
Pro β-CH2
Pro γ-CH2
Glu β-CH2
Arg β-CH2
Lys δ-CH2 + β-CH2
Arg γ-CH2
Leu β-CH2 + γ-CH
Lys γ-CH2
Ala -CH3
Thr -CH3
Val -CH3
Leu -CH3
Ile -CH3
3
2
1
0
PPM

(2) Histone I in Reconstituted DNA-histone I Complex

DNA-histone I complex for our present study was prepared from a solution containing 5.4 mg/ml of DNA and 2.5 mg/ml of histone I by continuous salt gradient dialysis. It is a low converage preparation with 1.2 amino acid residue of histone I per phosphate (a.a./p).

As shown in Fig. 2, significant changes of histone I in DNA-histone I complex were observed compared to that in the free state. All the spectral regions in the complexed state shift to low field in the order III (Ala CH_3, Lys γ CH_2) > I + II (aliphatic) > IV (Lys β CH_2 + δ CH_2) > VII (Lys ε CH_2) with differential chemical shift - 80 Hz, - 60 Hz, - 50 Hz, and - 10 Hz, respectively, at 18°C. These differential chemical shifts are defined as the chemical shifts relative to those of the free histone I under the same experimental conditions. Region V which is due to the resonances of β-CH_2 group of glutamic acid residues of this protein appears to be broad beyond detection in the complex. The apparent linewidth of the observed spectral regions does not change significantly in the complexed state as compared with that in the free histone I.

The temperature dependence of the differential chemical shifts of each spectral region of histone I in DNA-histone I complex is given in Fig. 3 . As the temperature is increased, the spectra of the complex qualitatively remain the same, but all the spectral regions shift progressively to high field, especially after 60°C. At 90°C, when the whole DNA in the complex is thermally denatured, the chemical shifts of most of

Figure 2. 220 MHz pmr spectra of DNA-histone I complex at various temperatures (concentration histone I 2.5 mg/ml DNA 5.4 mg/ml).

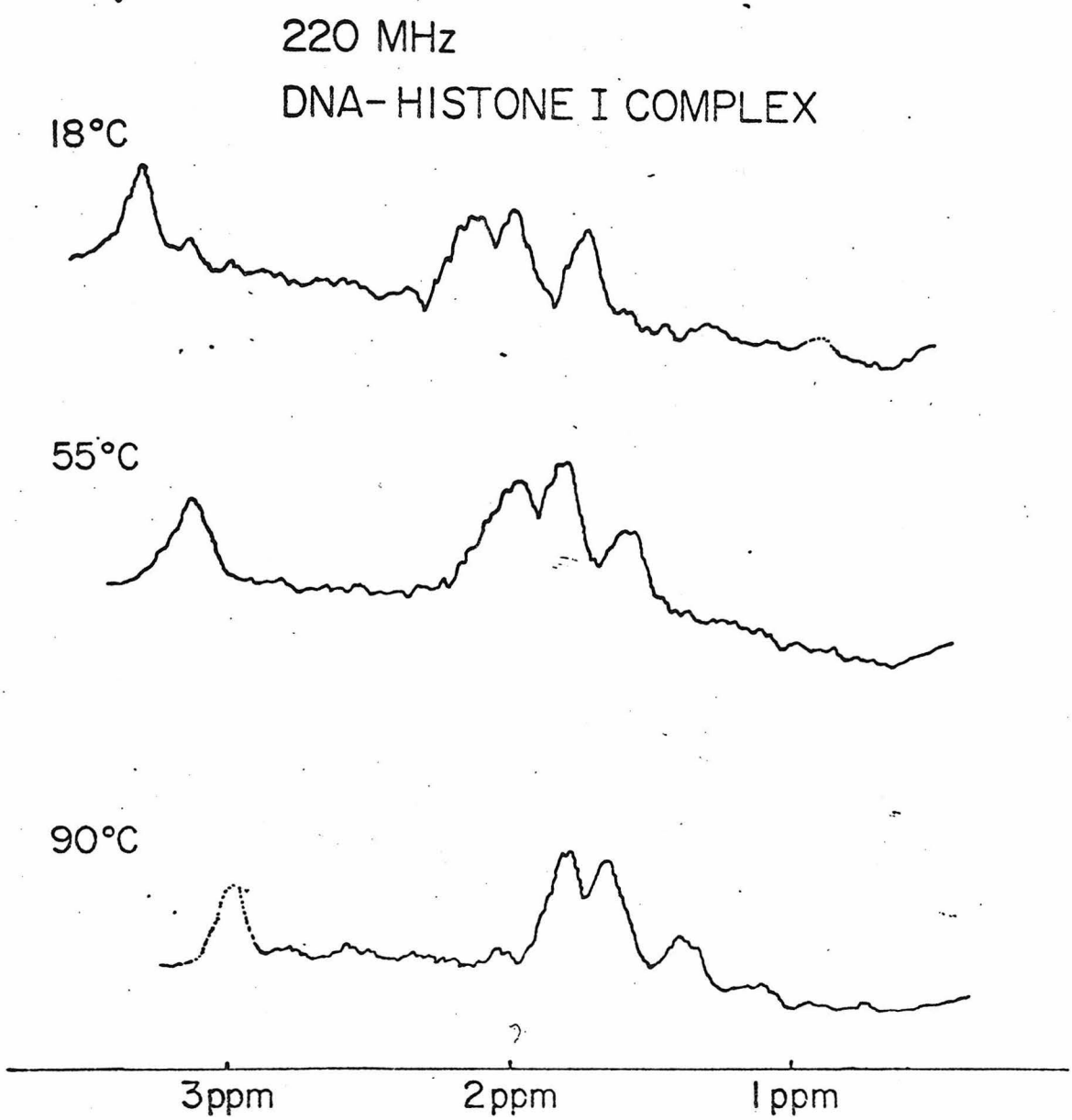
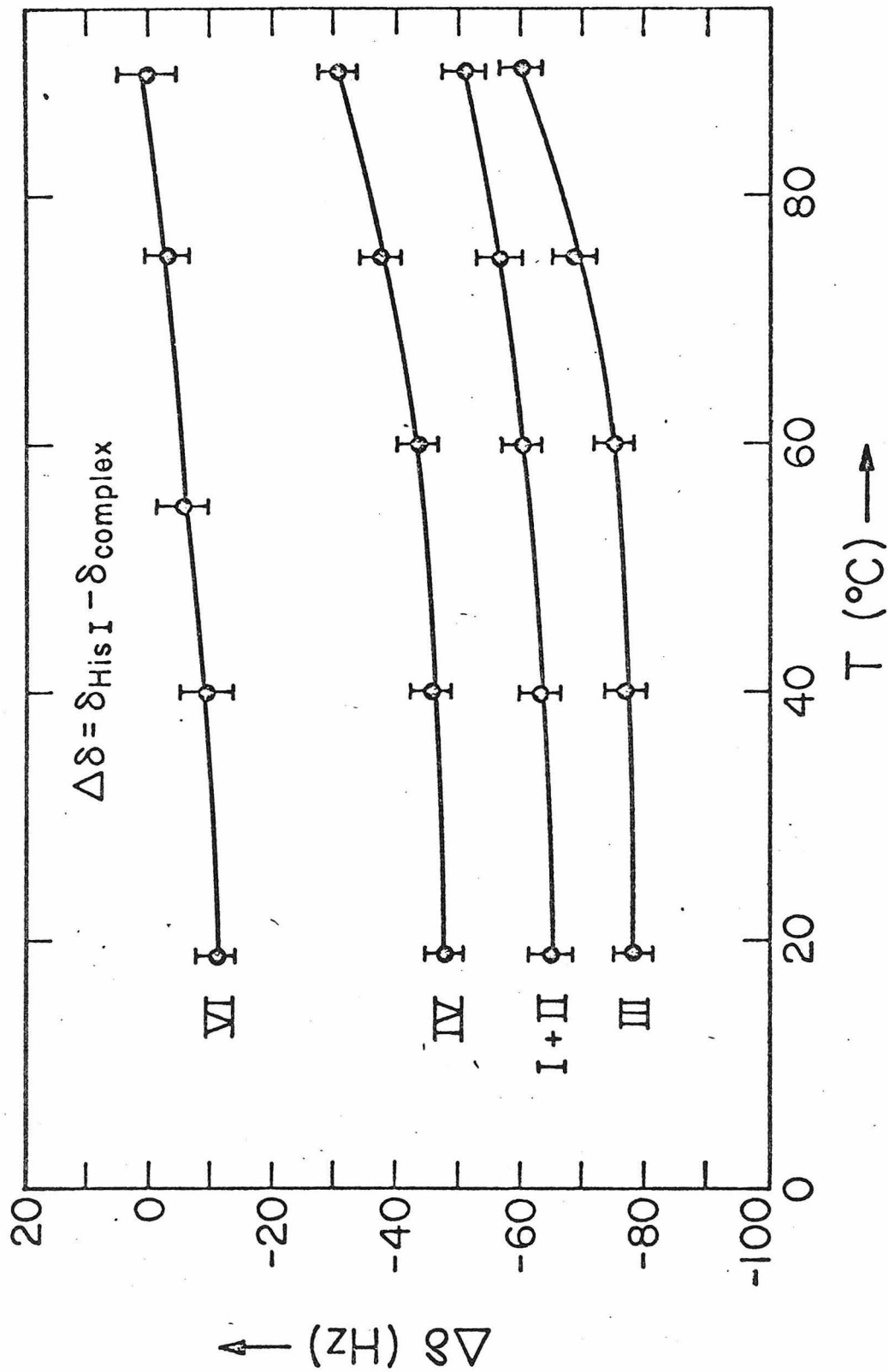

220 MHz
DNA-HISTONE I COMPLEX
18°C
55°C
90°C
3ppm
2ppm
1ppm

Figure 3. Differential chemical shift of each spectral region of histone I in DNA-histone I complex at various temperatures.


Δδ (Hz)
T (°C)
Δδ = δHis I − δcomplex
VI
IV
I+II
III
20
0
-20
-40
-60
-80
-100
0
20
40
60
80

the spectral regions still remain low field shift relative to those of histone I in the free state.

From the absolute area measurements, it was found that at room temperatures, the total area observed in the complex is only 15 ± 3% of that observed for the free histone I with the same histone concentration (2.5 mg/ml) at which the effect of histone aggregations is negligible. With respect to the spectral regions for basic amino acid side chains, only 13% spectral area was observed in the complex compared to that in the free state, while for those of the aliphatic amino acid residues, about 20% was observed. The spectral area measurements at different temperatures show that at 70°C, or higher, temperatures the total area observed from region I to IV increases only 20% compared to that observed at room temperature. (Fig. 4) So the area of spectral regions for the basic amino acid residues increases from 13% to 20% at high temperatures, while that for aliphatic amino acid regions remains essentially unchanged for the whole temperature range of interest. This is obviously indicated by the fact that the relative spectral area ratio between the spectral regions corresponding to the basic and the aliphatic amino acid residues, $[(A_{III+IV})/(A_{I+II})]$ increases from 1.7 at 18°C to 2.7 at 90°C. (Fig. 5). A sharp increase of this ratio was observed at temperatures around 60 to 70°C.

DNA-histone I complex of high coverage preparation with amino acid residue of histone I per phosphate of DNA (a.a./p) around 3.0 shows turbidity indicating higher aggregates. The quantitative study was limited and was not pursued.

Figure 4. Relative spectral area of histone I in DNA-histone I complex as a function of temperature.

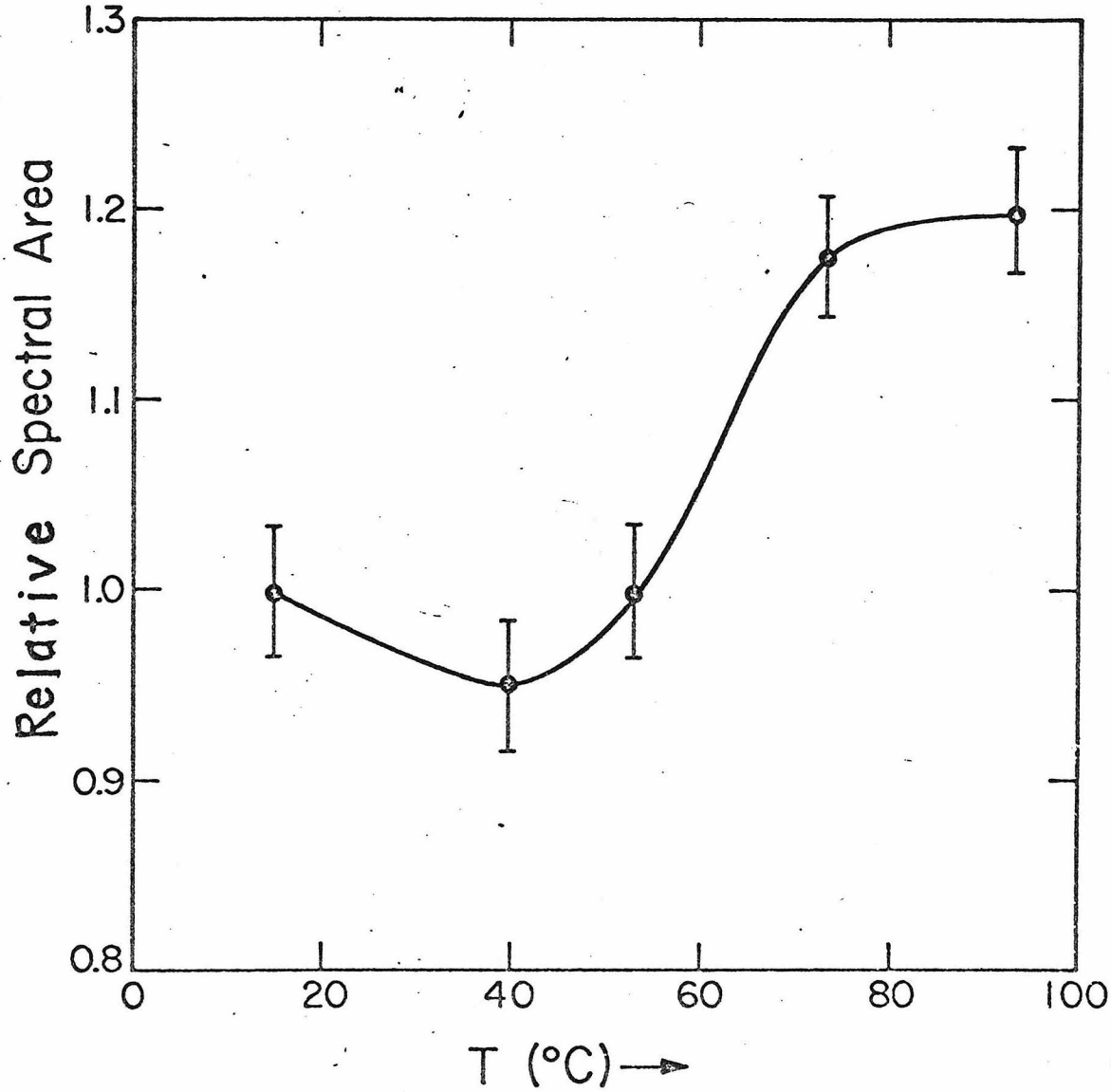
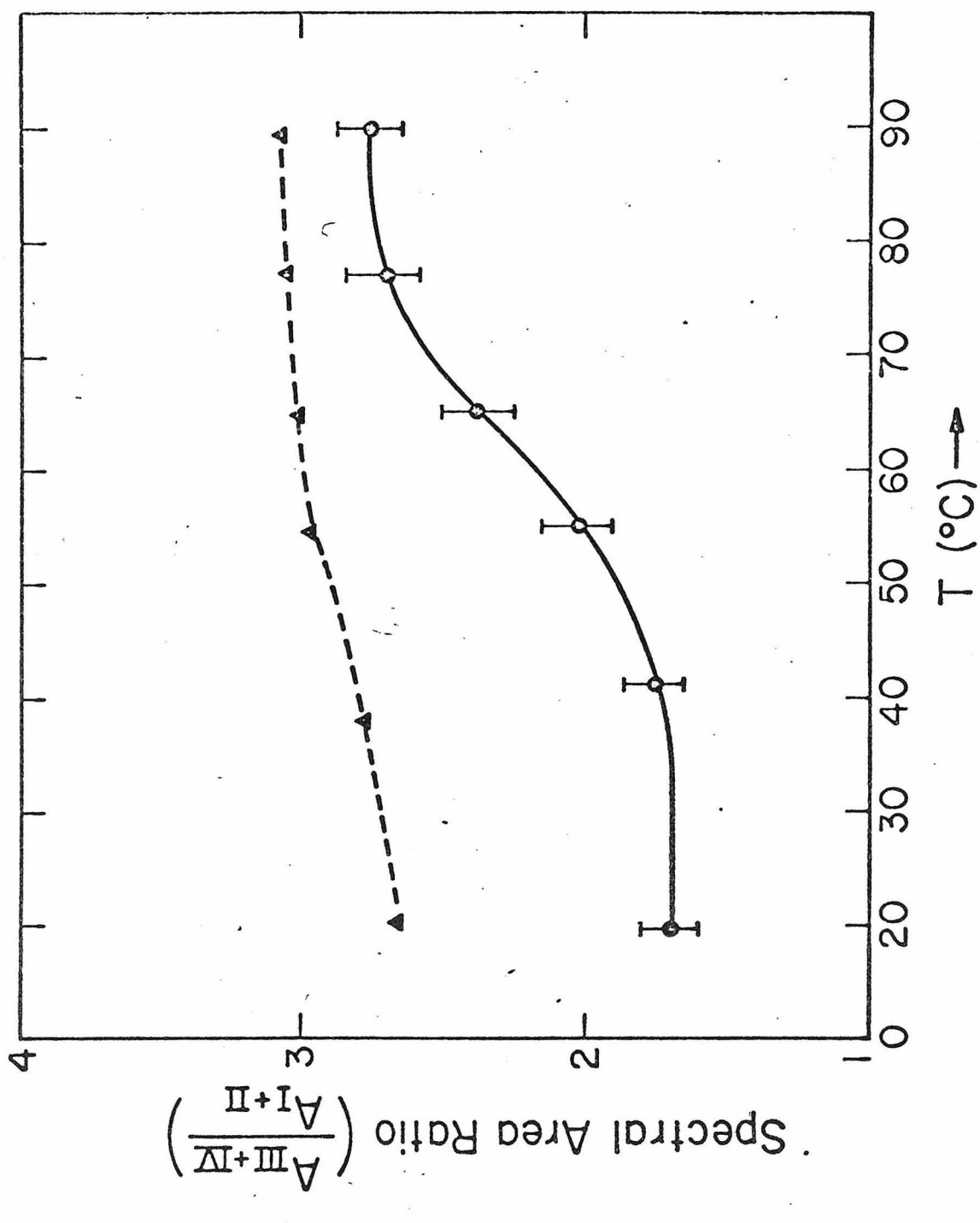

Relative Spectral Area
1.3
1.2
1.1
1.0
0.9
0.8
0
20
40
60
80
100
T (°C)

Figure 5. Area ratio of spectral region III + IV to I + II $[(A_{III+IV})/(A_{I+II})]$.


Spectral Area Ratio $\left(\frac{A_{III+IV}}{A_{I+II}}\right)$
4
3
2
1
0
20
30
40
50
60
70
80
90
T (°C) ⟶

DNA-histone I complex appears as white precipitate when pD is below 5.0. On the other hand, the solution is clear at high pD (pD > 10) and the histone I spectrum in the complex is similar to that of free histone, except that the apparent linewidth of each spectral region is still broader than those of free histones at the same pD.

The addition of high concentration of salt such as NaCl changes significantly the spectral area of the complex. But the resonances of each spectral region remain broad compared to those of free histone I at the same salt concentrations.

(3) Histone I when Mixed Directly with DNA or Polyuridylic Acid

We have also examined the spectral changes of histone I by direct mixing with DNA or polyuridylic acid. (Fig. 6)

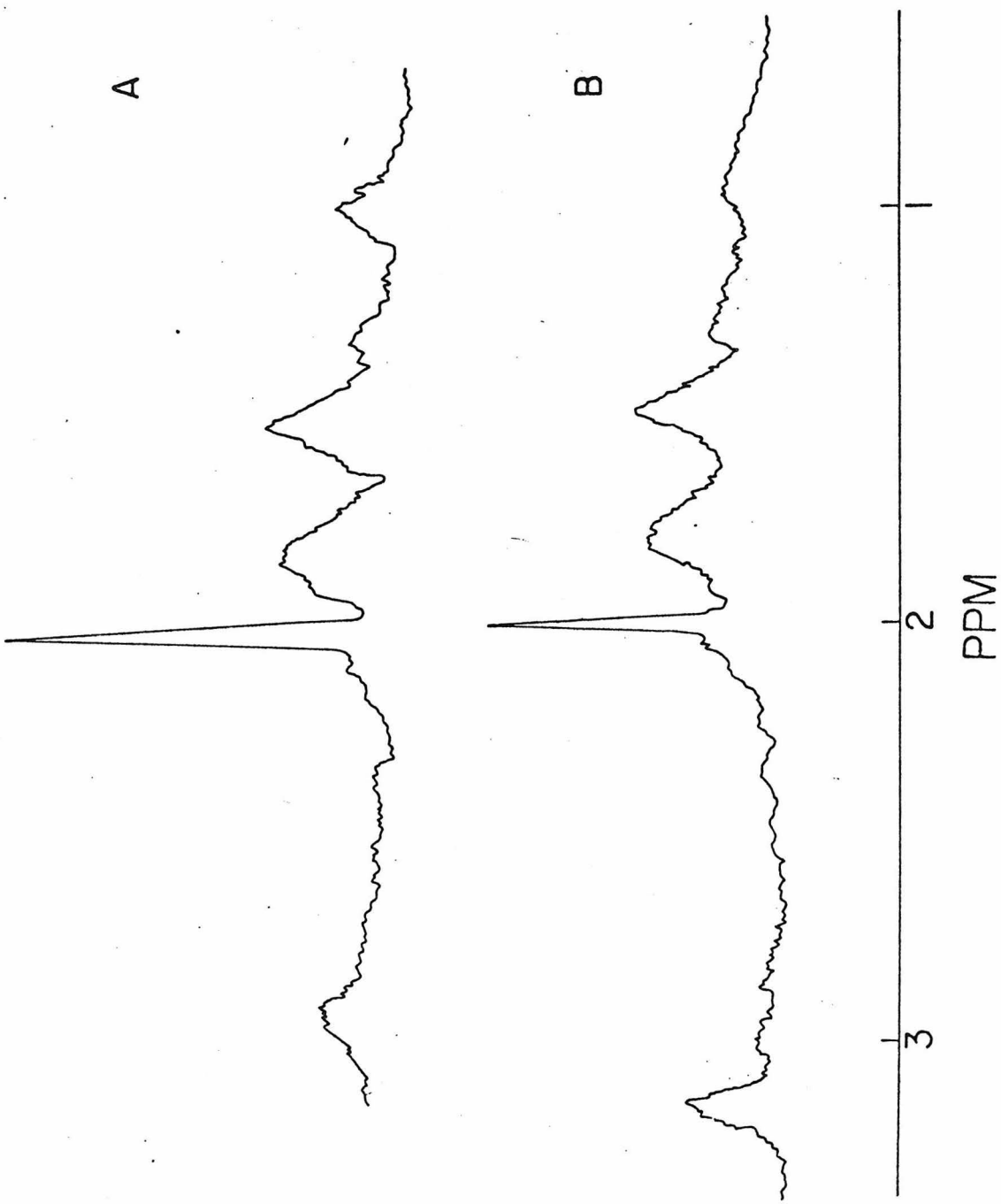
When histone I was mixed directly with DNA in aqueous solution, progressive changes of histone I spectrum were observed depending on time as well as the input ratio of histones to DNA (a.a./p) and the DNA concentrations. At a ratio of a.a./p = 0.7 and DNA concentration 3.2 mg/ml, the resonances of all the spectral regions of histone I are very broad beyond detection at room temperature. At a ratio of a.a./p = 1.8 and DNA concentration 3 mg/ml , all the spectral regions are observable in the pmr spectra but remain very broad. At a ratio of a.a./p = 4.5 and DNA concentration 1.6 mg/ml, the observed spectrum of histone I is qualitatively similar to that in the absence of DNA but all the spectral regions still remain somewhat broad, especially that corresponding to the

ε-CH_2 group of the lysine residues of this protein. At the initial stage of the last preparation only 40% spectral area was observed compared to that of the free histone I at the same concentration. About 10 to 30 Hz of downfield shift relative to that of the free histone I is observed for each spectral region when DNA is present in solution. The temperature study of the last preparation shows that the chemical shifts of each spectral region gradually approach those of the free histone I as the temperature is increased, and the observed spectral area appears to increase with increasing temperature.

When histone I is mixed directly with single-stranded polyuridylic acid, the resulting spectrum is similar to that of the free histone I for all the input ratios between histone I and poly U. However, some of the spectral regions of histone I appear to be broad and a significant area loss was observed when poly U is added to the solution. For example, at a ratio of a.a./p $\backsim$ 1.8 and poly U concentration of 5 mg/ml, only 60% of the spectral area was observed compared to that in the free state and each spectral region is about 10 to 20 Ha of low field shift relative to that of free histone I. Significant broadening and area loss were observed in the spectral region I corresponding to the aliphatic amino acid side chains. The broadening of the H_6 resonance on uracil bases of poly U was also noted.

As the temperature is increased, all the spectral region shift progressively to the high field and the apparent spectral width becomes narrow with increasing temperature. In this study, the spectral area loss

Figure 6. 220 MHz histone I spectra when mixing directly with DNA (A) or polyuridylic acid (B).
(Concentration: A. histone I 3 mg/ml, DNA 2 mg/ml)
B. histone I 3 mg/ml, poly U 5 mg/ml)).


A
B
3
2
1
PPM

of histone I seems to increase with decreasing a.a./p ratio when either DNA or polynucleotide is added to the histone I solution. However DNA is much more effective to cause the area loss of histone I than that of poly U.

4. DISCUSSION

Histone I in neutral aqueous solution is essentially a random coil polypeptide and processes well defined pmr spectral regions corresponding to the resonances arising from the side chain protons of various amino acid residues in this protein. However, when it is complexed with DNA, only several of the spectral regions are observed, i.e., regions I and II, III, IV, and VII. Compared to histone I in the free state, each observed spectral region exhibits -10 to -80 Hz of downfield differential shift. Region I and II corresponding to the side chain protons of the aliphatic amino acid residues collapse into one region. Regions V and VI are broad beyond detection in the complexed state. Spectral area measurements show that the spectral area observed for histone I in the DNA-histone I complex is only 15% of that observed in free histone I at the same concentration. All the evidences seem to indicate that the transition of histone I from the random coil to the ordered or rigid structure has occurred when it is complexed with high molecular weight DNA.

From the absolute area measurements of each spectral region at various temperatures, we found that percentage of the total side chain protons of lysine residues observed in the complex increases from 13% to 20% at temperatures where the whole DNA is thermally denatured. While those of the aliphatic amino acid residues remain unchanged for the whole temperature range. This is also indicated by the increase of the area ratio $[(A_{III+IV})/(A_{I+II})]$ from 1.7 to 2.7 around 60 to 70°C where the denaturation of DNA in the histone I covered region is known to occur. The slight and sharp increase of the relative spectral area in the region of lysine side chains indicates the partial dissociation of these residues from DNA and/or the increase in motional freedom of the side chain of these residues due to the DNA denaturation.

It is not unreasonable to assume that the histone I resonances in the DNA-histone I complex are too broad to be observed if they arise from those amino acid residues which are tightly bound or neighboring to those tightly bound to the rigid DNA. Hence the resonances observed in the pmr spectra of this complex would be from those of weak binding and/or free from binding to DNA. The fact that the apparent linewidth of each observed spectral region of histone I in the complex does not differ much from that of the free histone I and is not sensitive to temperature changes seems to support this suggestion. Since both the observed spectral area and the chemical shifts of most of the spectral regions do not go back to those of the free histone I at high temperatures, we may conclude that histone I is still not removed from DNA when the latter is thermally denatured.

We have studied in detail about the amino acid distribution and the binding nature with DNA of this protein to account for 15% of the histone I spectral area observed in the complexed state. Since the primary sequence of this protein is still unknown, the information can only be obtained from the analysis of its amino acid distributions. From this analysis, the whole polypeptide chain of a histone I molecule can be divided into three regions according to their characteristic amino acid distributions. The first region is the carboxyl terminal half of this protein containing 116 amino acid residues (residues 101 to 216). This region consists of about 30% of the total aliphatic amino acid residues (including Leu, Ileu, Val and Thr), 75% of the total lysine residues, and 20% of the total acid amino acid residues of this protein. The ratio of the basic to acidic amino acid residues is as high as 15. The amino terminal half of this protein includes two polypeptide regions, i.e., the second region (residues 52 to 100) and the third region (residues 1 to 51), respectively. The percentages of the total aliphatic amino acid residues in these two regions are 50% and 20% respectively. While the basic and acidic amino acid residues are almost evenly distributed in these two regions with basic to acidic amino acid ratio of 1.1 to 1.3 respectively.

From this analysis, we may regard the carboxyl-terminal half and the amino-terminal half of this protein as the strong and weak binding regions to DNA respectively, since their basic to acidic amino acid ratios and the charge distributions are known to be quite different. This classification is based on the generally accepted assumption that the

electrostatic interactions between the phosphate group of DNA and the $\varepsilon\ NH_3^+$ group of lysine residues are essential for the complex formation. In the N-terminal half of this protein, the low percentage of lysine residues and the comparable amount of acidic amino acid residues weaken considerably the binding of this part of the protein with DNA and make this half of histone I molecule in the complex exhibit higher degree of motional freedom. If this is the real situation, the fraction of the resonances observed in the complex could predominantly arise from N-terminal half of this protein.

However, the observed percentage of the spectral area of histone I in this complex does not quite agree with that expected from the corresponding amino acid residues in this half of the protein assuming that they are in a relatively free state. According to our estimations, only 50% of the lysine residues and 30% of the aliphatic amino acid residues in this half of the protein were observed. Further classification of this half of histone I into two equal regions is necessary for the interpretation of our experimental observations. The second region (amino acid residues 51 - 100) of the peptide chain is adjacent to the strong binding region of this protein. The restricted and anisotropic motions of the side chains of the amino acid residues of this region is expected to occur. Hence the resonances arising from this region are possibly too broad to be observed. While the third region (a.a. residues 1 - 51) of histone I molecule is farther away from the strong binding region and the motion would be relatively free and less restricted. Most

of the resonances arising from this region is expected to be observable in the pmr spectrum of this complex. In fact, our experimental observations are not exactly inconsistent with this interpretation. The observed spectral area of histone I in the complex is almost equal to that estimated from the side chain protons of both the basic and aliphatic amino acid residues of this particular region of the protein assuming that they are in a relatively free state.

Besides the electrostatic interaction, the hydrophobic interaction between the hydrophobic side chains of some aliphatic amino acid residues and DNA backbone is noted to be important for the complex formation. This is clearly indicated by the observation of an equally significant area loss of the side chain protons of both aliphatic amino acid residues and lysine residues. This evidence is consistent with the result of Olin's solvent dependent study of the melting temperature of DNA-histone I complex,[9] for the latter also showed that the hydrophobic interaction between DNA and histone I is also important in the stabilization of the complex.

The observed significant downfield shift of each spectral region in the complexed state would be predominantly due to the specific interaction of histone I with DNA, since the polyelectrolyte nature of DNA changes the solvent structure and provides a specific ionic environment to the tightly bound histone I. The contribution from the anisotropic chemical shift due to the anisotropic orientation of the DNA chain is also possible to give significant downfield shift to the resonances

of histone I in the complex. Besides these, the minor contribution from the ring current effect of DNA bases inducing downfield shift to histone I resonances should also be considered, if they sit around the DNA grooves with 3 to 5 Å in distance.

The appearance of spectral region VII arising from the resonance of terminal ε-CH_2 group of lysine residues at room temperature would indicate that some of the lysine residues are in slow exchange with DNA phosphate group or free from interaction with DNA. While at high temperatures (> 50°C), this region becomes broad and difficult to assign, this could be due to weight averaged fast exchange with DNA or due to the interference from the sidebands of solvent (DHO) and Tris buffer in solution.

As evidence for the irreversibility of the complex, when heating the sample to 90°C and then cooling back to room temperature, the spectrum of histone I in the renatured complex looks different from the original one, both in chemical shift and linewidth. This would indicate that after DNA denaturation at high temperatures, the state of the renatured complex is different from the original state. In the presence of the salt, the increase in the spectral area and the relative area of lysine residues would imply the partial dissociation of histone I from DNA. The recovery of histone I spectrum at high pD's also indicates the dissociation of histone I from DNA.

In order to show the characteristic structure of this DNA-histone I complex prepared by salt gradient dialysis, we study the spectral

behavior of histone I by mixing it directly with native DNA or single-stranded poly U as a controlled experiment for comparison. In the presence of DNA or poly U, histone I spectra become broad and exhibit a significant area loss compared to those in the free state. At the same weight concentration, DNA is more effective to induce area loss of histone I spectra compared to poly U.

It is known that both DNA or poly U are polyelectrolyte in nature. The presence of these two molecules in solution will induce strong electrostatic interactions between the phosphate group of these two molecules and ε-NH_3^+ group of lysine residues of histone I. These interactions would result in (1) the complex formation between histone I and DNA or poly U, and/or (2) the possible self-aggregation of histone I molecules. Because of the high molecular weight and the rigidity of native DNA, it behaves like a rigid rod in solution. (4×10^6) The complex is more favorable when it is mixed with histone I in solution. A significant spectral area loss of histone I, especially in the region corresponding to the ε-CH_2 group of lysine residues at the beginning of the preparation indicates that the exchange rate of histone I between the free and the bound state is slow compared to the inverse of the transverse relaxation times when histone I is bound to DNA. The time dependent area loss of histone I spectra in this mixture would also imply that the formation of a more ordered structure in the "complex" is a very slow process.

Poly U is a random coil polyelectrolyte with a lower molecular weight ($\sim 10^5$) compared to that of DNA. Its interaction with histone I

is essentially random and electrostatic in nature. It is expected that there is little possibility to form a rigid and ordered complex with histone I. On the other hand, a significant area loss of histone I in the regions corresponding to the aliphatic amino acid residues is observed when poly U is added to the histone I solution. This would suggest the possibility of the hydrophobic aggregation of histone I due to the presence of poly U in solution. The magnitude and the temperature dependence of the differential chemical shifts of each spectral region of histone I in the direct-mixing preparation are very different from those of the reconstituted DNA-histone I complex. This also strongly implies that the structures of histone I in the solution prepared by direct mixing with DNA or poly U are not the same as those of the reconstituted complex prepared by continuous salt gradient dialysis.

5. BIOLOGICAL IMPLICATION OF THE DNA-HISTONE I RECONSTITUTED COMPLEX

Besides the high content of the basic amino acid residues, histones are also characterized by their assymmetric distributions of the charged amino acid residues along the peptide chains of the proteins. Usually one half of a histone molecule is enriched in the basic amino acid residues and is presumably the strong binding site to DNA. While another half consists of almost a comparable number of the basic and acidic amino acid residues and resembles a small nonhistone protein which might

have specificity for factors other than DNA. The results of our study on the DNA-histone I complex also suggest that part of the peptide chain in the amino terminal half of the protein is relatively free in the complex state.

Recent studies on the transcription of DNA by RNA polymerase indicated that the percentage of RNA transcribed from DNA of a reconstituted nucleohistone is actually more than that indicated by the basic amino acid to phosphate ratio.[19] This suggests that part of the basic amino acid residues is not actually tightly bound to DNA in the reconstituted nucleohistone so that it does not inhibit the transcription of DNA by RNA polymerase. A recent study about the structure of chromatin by Clark et al.[20] indicated that almost half of DNA in chromatin is not covered with nucleoproteins. This would be somewhat correlated to the structure of the nucleohistones, i.e., half of the polypeptide chain of histones is loosely or not bound to DNA at all.

Our pmr study of reconstituted DNA-histone I complex has suggested a structural model of nucleohistone I, i.e., about one-quarter of the polypeptide chain of the histone I molecule is free from binding to DNA and, therefore, can be observed in the pmr spectra.

REFERENCES

1. R. C. C. Huang and J. Bonner, Proc. Nat. Acad. Sci. U.S., 48,1216 (1962).
2. G. C. Barr and J. A. V. Butler, Nature 199,1170 (1963).
3. V. G. Allfrey, V. C. Littau, and A. E. Mirsky, Proc. Nat. Acad. Sci. U. S., 49, 414(1963).
4. E. O. Akinrimisi, J. Bonner, and P. O. P. Ts'o, J. Mol. Biol., 11, 128(1965).
5. M. Tsuboi, K. Matsuo, and P. O. P. Ts'o, J. Mol. Biol., 15,256(1966).
6. M. Leng and G. Felsenfeld, Proc. Nat. Acad. Sci. U.S., 56,1325(1966).
7. D. E. Olins, A. L. Olins, and P. H. von Hippel, J. Mol. Biol., 24, 157(1967).
8. D. E. Olins, A. L. Olins, and P. H. von Hippel, J. Mol. Biol., 33, 265(1968).
9. D. E. Olins, J. Mol. Biol., 43,439(1969).
10. T. Shih, and J. Bonner, J. Mol. Biol., 48, 469(1970); also, 50,333 (1970).
11. H. J. Li and J. Bonner, Biochemistry, 10,1461(1971).
12. M. Boublik, E. M. Bradbury, and C. Crane-Robinson, Eur. J. Biochem, 14,486(1970).
13. M. Boublik, E. M. Bradbury, C. Crane-Robinson, and E. W. Johns, Eur. J. Biochem., 17,151(1970).

14. M. Boublik, E. M. Bradbury, C. Crane-Robinson, and H.W.E. Rattle, Nature, 229,149(1971).

15. E. W. Johns, Biochem, J. 2,55(1964).

16. R. C. C. Huang, J. Bonner, and H. Murray, J. Mol. Biol., 8,54(1964).

17. C. C. McDonald and W. D. Phillips, J. Am. Chem. Soc. 91,1513(1969).

18. G. C. Y. Lee, J. H. Li, and S. I. Chan, Biochemistry, in press (1972).

19. T. Shih, Ph.D. Thesis, California Institute of Technology (1969).

20. R. J. Clark and G. Felsenfeld, Nature, 229,101(1971).

PART III

NUCLEAR MAGNETIC RESONANCE STUDIES OF RIBONUCLEASE A- INHIBITOR INTERACTIONS AND NUCLEOTIDE CONFORMATIONS

1. A ^{31}P NMR STUDY OF THE ASSOCIATION OF URIDINE-3'-MONOPHOSPHATE TO RIBONUCLEASE A

Nuclear magnetic resonance spectroscopy (nmr) is now extensively used to probe the active sites of enzymes in solution.[1,2] Most of these studies have been concerned with enzyme-inhibitor interactions[3-7] and information about the active site is often inferred from the effects which the enzyme has on the chemical shifts,[8,9] linewidths,[10] as well as the relaxation times[11,12] of specific resonances of the inhibitor. Of the various types of nmr experiments, high resolution proton magnetic resonance spectroscopy (pmr) has been the most widely employed. While it is often possible to infer structural information from these studies, it is in general difficult to obtain kinetic information because of small chemical shifts and the resultant long time scales of observation.

In this chapter, we wish to report a ^{31}P nmr study of the association of ribonuclease A (RNase A) with its product inhibitor uridine-3'-monophosphate (3'-UMP). Because ^{31}P chemical shifts are usually large, the nmr observation time is, in general, comparable to the time scales of the enzyme-inhibitor interactions, even at relatively low enzyme concentrations. Moreover, the solvent in these experiments can be H_2O instead of D_2O, so one is no longer faced with questions often raised in connection

with the pmr studies regarding the isotopic effect of the solvent on the protein conformation and the enzyme-inhibitor interactions under consideration.

The RNase A used in this work was isolated from bovine pancreas and was obtained from Calbiochem, Los Angeles, California, as an A grade reagent. Its biological activity was shown to be 65 Kunitz units/mg.[13] A grade 3'-UMP was purchased from Calbiochem in the acid form, and was converted to the sodium salt on a Dowex 50W-X8 cation exchanger.

Solutions containing 0.1 M of the inhibitor and various concentrations of the enzyme (1×10^{-4}M to 2×10^{-3}M) were prepared in H_2O, and their pH was maintained at 7.5 ± 0.01 using 0.1 M tris buffer and 0.05 M NaCl. The ^{31}P nmr experiments were performed at a magnetic field of 53 kgauss and the resonance frequency of 89 MHz on a modified Varian HR-220 superconducting nmr spectrometer equipped with frequency sweep and multinuclei capabilities. Proton noise decoupling was used to remove all phosphorous and proton spin-spin couplings and a C-1024 time averaging computer was used to enhance the signal to noise ratio. The ^{31}P chemical shifts were measured relative to an external reference of trimethylphosphite sealed in a capillary. Unless specified otherwise, all the experiments were done at the normal probe temperature of 32 ± 2°C. 5 mm O.D. sample tubes were employed.

The effect of RNase A at various enzyme concentrations on the ^{31}P resonance of 0.1 M 3'-UMP is shown in Fig. 1. The ^{31}P resonance was observed to shift progressively to higher field and to become broader with increasing enzyme concentration. These chemical shifts and linewidths are summarized in Fig. 2 as a function of the enzyme/inhibitor concentration ratio. The observed monotomic variation of the chemical shift and linewidth with enzyme concentration suggests that the kinetics of the enzyme-inhibitor binding is either intermediate or rapid compared to the reciprocal of the ^{31}P chemical shift difference of the inhibitor between the complexed and the unbound states. Although the observed broadening may reflect the shorter transverse relaxation time of phosphorous nucleus when it is bound to the enzyme, an estimation of this transverse relaxation ($\sim$ 0.02 sec) reveals that it is too long to account for the observed linewidth at the low concentrations of the enzyme used. We therefore surmised that the observed line broadening arises predominantly from the intermediate chemical exchange of the inhibitor between the free and the bound states. In order to substantiate this point, we have simulated the ^{31}P spectrum on the digital computer for different enzyme/inhibitor concentration ratios over a range of chemical shifts of the inhibitor between the free and bound states (Δ) and as a function of the kinetic parameters for the overall enzyme-inhibitor reaction:

Fig. 1. ^{31}P spectra of 0.1 M 3'UMP under three different enzyme/inhibitor concentration ratios: number of times averaging at: (1) $[E_0]/[I_0] = 0$, 5 scans; (2) $[E_0]/[I_0] = 0.009$, 10 scans; (3) $[E_0]/[I_0] = 0.016$, 25 scans.

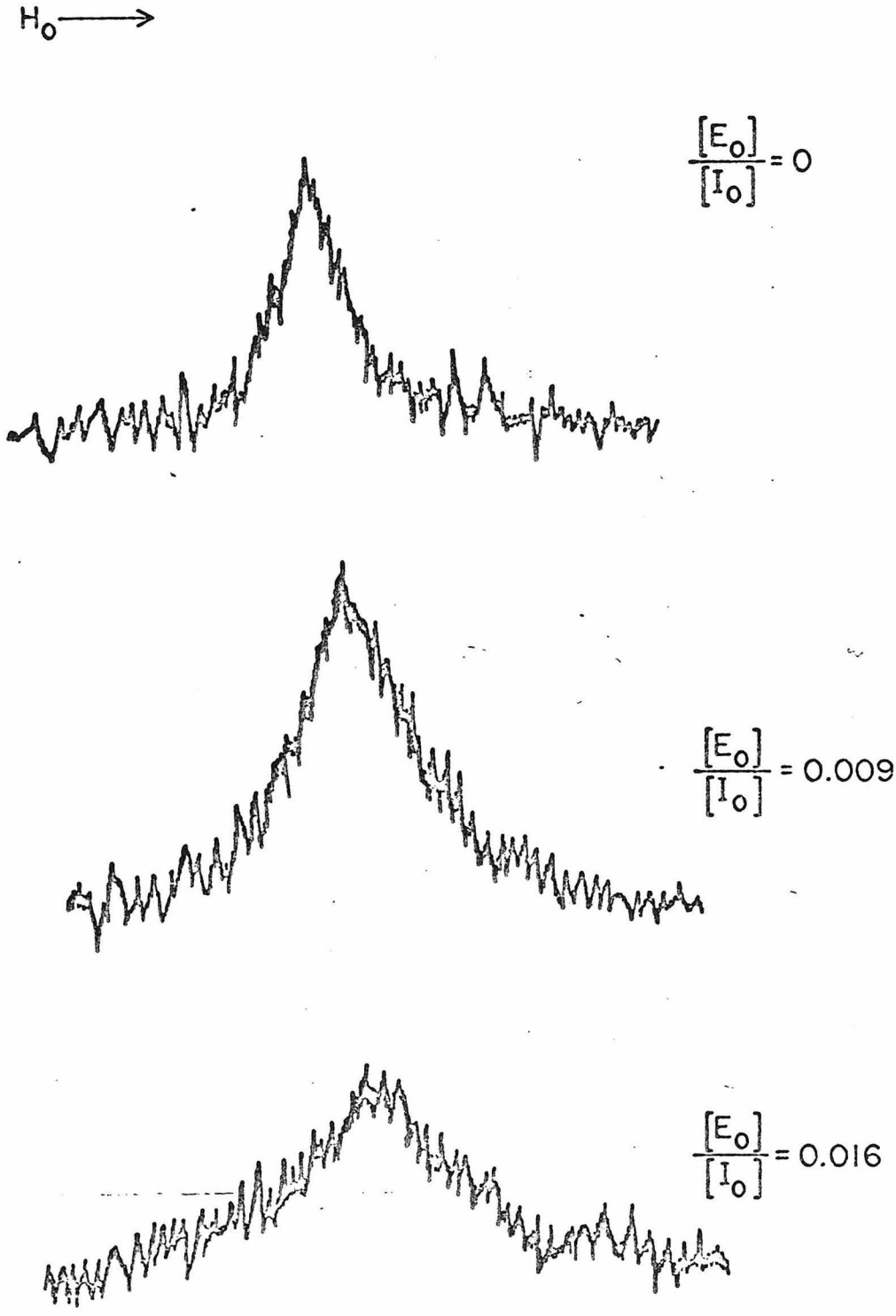
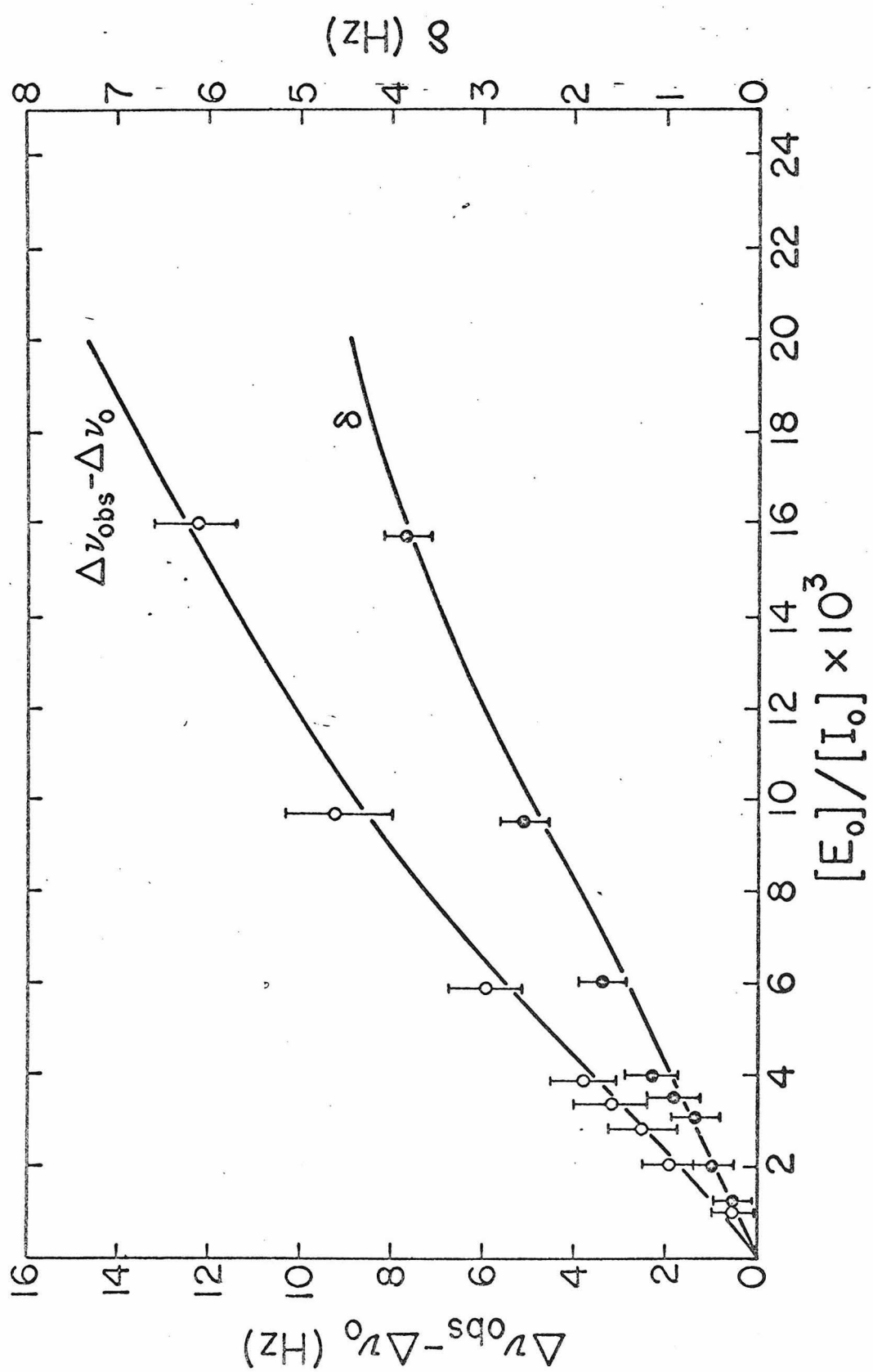

H_0
$\frac{[E_0]}{[I_0]} = 0$
$\frac{[E_0]}{[I_0]} = 0.009$
$\frac{[E_0]}{[I_0]} = 0.016$
25 Hz

Fig. 2. The chemical shifts (⊣●⊢) and the linewidth (⊣○⊢) ^{31}P resonance of 3′UMP under various enzyme/inhibitor concentration ratios. Where $\Delta\nu_{obs} - \Delta\nu_0$ is the difference in linewidth between the absence and the presence of the enzyme, δ is the observed chemical shift of the inhibitor in the presence of the enzyme relative to the free inhibitor. Solid curves are those simulated by the digital computer.


$\Delta\nu_{obs}-\Delta\nu_o$ (Hz)
δ (Hz)
$[E_o]/[I_o] \times 10^3$
$\Delta\nu_{obs}-\Delta\nu_o$
δ

$$E + I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\,I$$

E, I and EI denote the enzyme, inhibitor and enzyme-inhibitor complex, respectively, and k_1, k_{-1} are the association and dissociation rate constants. Although it is known that the interaction between RNase A and its inhibitors involves a multistep equilibrium, our nmr measurements are unfortunately only sensitive to those steps with rates comparable to the nmr time scale.

Computer simulation of ^{31}P spectrum clearly indicated that the ^{31}P broadening of the inhibitor arises from the intermediate chemical exchange of the inhibitor between its free and complexed states. In our analysis we have assumed that $[EI]/[I] \simeq [E_o]/[I_o]$, so that $k_1[E] = k_{-1}[EI]/[I] \simeq k_{-1}[E_o]/[I_o]$. This should be an extremely good approximation at high inhibitor concentrations and at low $[E_o]/[I_o]$ ratios. The equilibrium constant for the association process is $\geq 10^3\ M^{-1}\ell$. Hence at an inhibitor concentration of 0.1 M, $[E]/[EI] \leq 10^{-2}$, and the enzyme should be totally complexed for $[E_o]/[I_o]$ ratios of 10^{-3} to 2×10^{-2}.

Plots of the calculated chemical shifts and linewidths of the ^{31}P resonance versus k_{-1} for $\Delta \simeq 800$ Hz are presented in Figs. 3 and 4. Since the first order dissociation rate constant is independent of the enzyme or the enzyme/inhibitor concentration ratio at a given temperature, the Δ which characterizes the system must be that

Fig. 3. The computed chemical shift of ^{31}P resonance of 3′UMP as a function of the dissociation rate constant (k_{-1}) under various enzyme/inhibitor concentration ratios, assuming that the chemical shift difference (Δ) between the free and the bound states is 800 Hz.

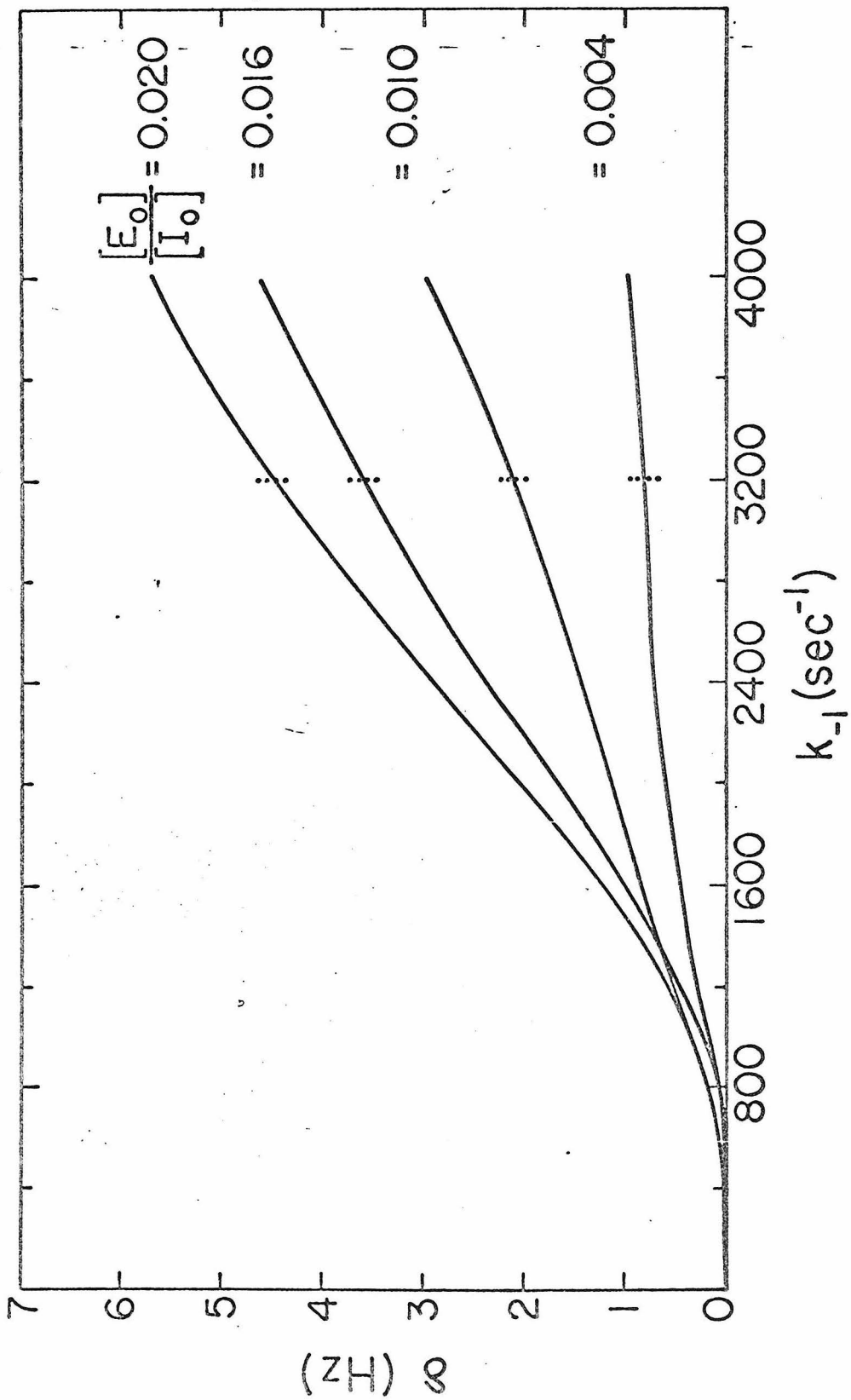
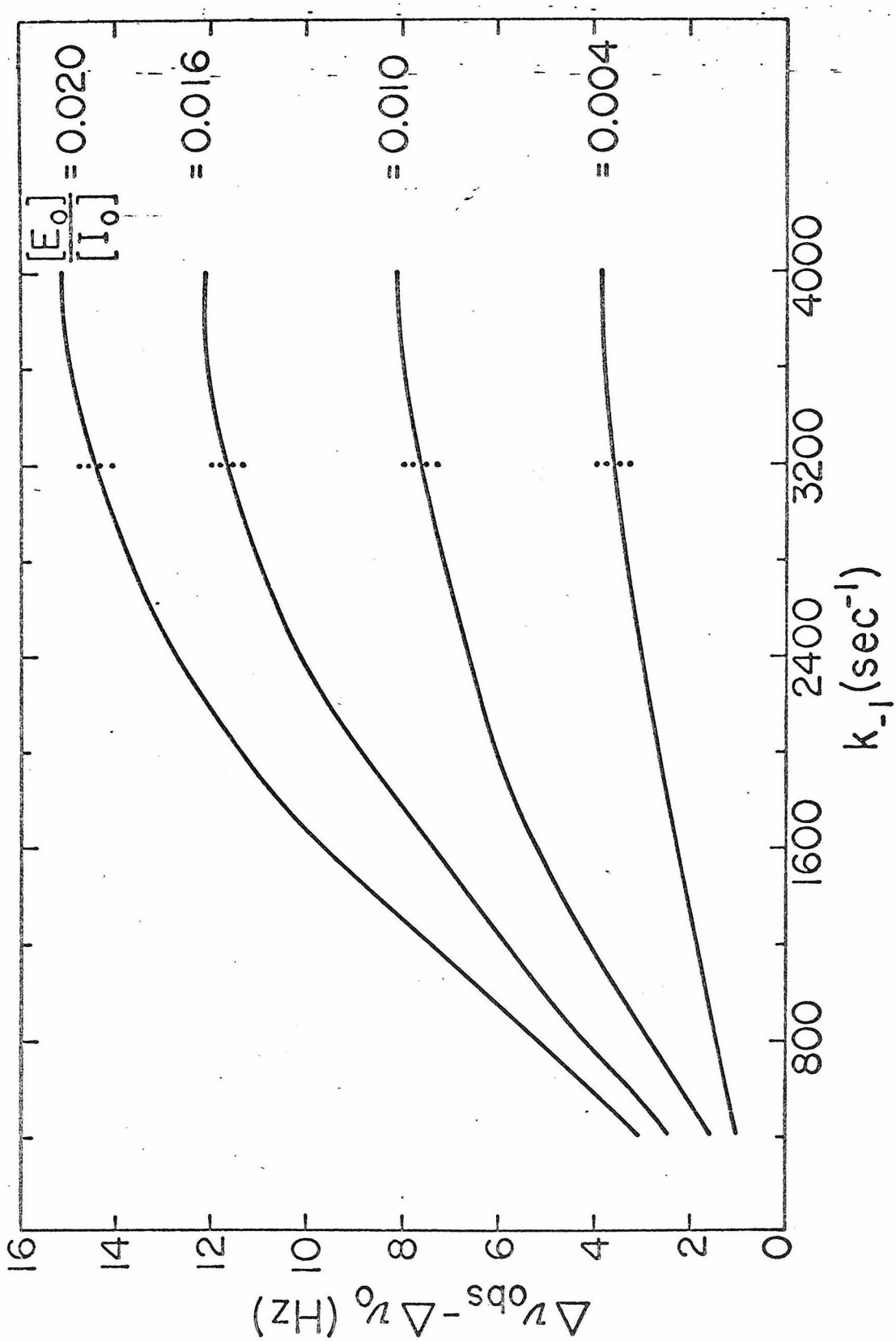

δ (Hz)
0
1
2
3
4
5
6
7
800
1600
2400
3200
4000
k_{-1} (sec^{-1})
$\frac{[E_o]}{[I_o]}$ = 0.020
= 0.016
= 0.010
= 0.004

Fig. 4. The computed linewidth difference ($\Delta\nu_{obs} - \Delta\nu_0$) of ^{31}P resonance of 3′UMP as a function of the dissociation rate constant (k_{-1}) under various enzyme/inhibitor concentration ratios, assuming Δ is 800 Hz.


$\frac{[E_o]}{[I_o]}$ = 0.020
= 0.016
= 0.010
= 0.004
$\Delta\nu_{obs} - \Delta\nu_o$ (Hz)
k_{-1} (sec^{-1})
0
2
4
6
8
10
12
14
16
800
1600
2400
3200
4000

yields the observed linewidths and the observed chemical shifts for the same value of k_{-1}. The best fit to our data was obtained for $\Delta \approx +800 \pm 50$Hz and $k_{-1} \sim 3200 \pm 300$ sec^{-1} at 32°C. The agreement between the observed and the calculated chemical shifts and the linewidths for various enzyme/inhibitor ratios is depicted in Fig. 2 and can be seen to be excellent.

The activation energy for the dissociation process can be determined from a temperature study of ^{31}P resonance widths at fixed enzyme/inhibitor ratios. Examination of the linewidth changes for three enzyme/inhibitor ratios over the temperature range of 10° to 30°C yielded an energy of activation of 3.5 ± 1.0 kcal/mole.

The kinetic parameters which we have deduced from the present ^{31}P nmr study of the association of 3'-UMP to RNase A are in general agreement with the temperature jump measurements of Hammes *et al.*,[14] who succeeded in resolving the kinetics of each step of the enzyme-inhibitor association. Jardetzky *et al.*,[7] have suggested, on the basis of their pmr studies of RNase A-nucleotide complexes and the x-ray structure of RNase S[15,16] completed by the Wyckoff group at Yale, that the phosphate group of cytidine-3'-monophosphate is simultaneously bound in the dianionic form to the positively charged side-chains of two amino acid residues, lysine 41 and histidine 119, at the active site of the enzyme. Although similar pmr work on 3'-UMP has

not been reported, one would nevertheless expect similar bonding of the phosphate group at the active site for this inhibitor. If this were the case, we would expect the ^{31}P resonance of the inhibitor to shift upfield by $\sim$ 800 Hz at 89 MHz upon binding to the enzyme at a solution pH of $\sim$ 7.5, where the phosphate exists as the dianion prior to association to the enzyme. Preliminary ^{31}P studies of nucleotides undertaken in this laboratory have shown that protonation of a primary phosphate, in general, shifts the ^{31}P resonance upfield by $\sim$ 400 Hz (at 89 MHz) per protonation step. Thus, the ^{31}P chemical shift of the inhibitor upon association can help to elucidate the involvement of the phosphate group in the enzyme-inhibitor interaction.

References

1. O. Jardetzky, N.G. Wade, and J.J. Fisher, Nature, 197, 183 (1963).

2. D.P. Hollis, Biochemistry, 6, 2080 (1967).

3. E.W. Thomas, Biochem. Biophys. Res. Commun., 24, 611 (1966).

4. T. McL. Spotswood, J.M. Evans and J.H. Richards, J. Amer. Chem. Soc., 89, 5052 (1967).

5. D.H. Meadows and O. Jardetzky, Proc. Nat. Acad. Sci., 61, 406 (1968).

6. D.H. Meadows, O., Jardetzky, R.M. Epand, H.H. Ruterjans, and H.A. Scheraga, Proc. Nat. Acad. Sci., 60, 766 (1968).

7. D.H. Meadows, G.C.K. Roberts and O. Jardetzky, J. Mol. Biol. 45, 491 (1969).

8. M.A. Raftery, F.W. Dahlquist, S.I. Chan and S.M. Parsons, J. Biol. Chem. 243, 4175 (1968).

9. F.W. Dahlquist and M.A. Raftery, Biochemistry 7, 3269 (1968).

10. J.T. Gerig and J.D. Reinheimer, J. Amer. Chem. Soc., 92, 3146 (1970).

11. B.D. Sykes, J. Amer. Chem. Soc., 91, 949 (1969).

12. B.D. Sykes, P.G. Schmidt and G.R. Start, J. Biol. Chem. 245, 1180 (1970).

13. M. Kunitz, J. Biol. Chem. 164, 563 (1946).

14. G.G. Hammes and F.G.Walz, Amer. Chem. Soc., 91, 7179 (1969).

15. G.B. Kartha and D. Harker, Nature 213, 862 (1967).

16. H.W. Wyckoff, K.D. Hardman, N.M. Allewell , T. Inagemi, L.N. Johnson and F.M. Richards, J. Biol. Chem., 242, 2398 (1967).

2. SOLVENT EFFECTS ON NUCLEOTIDE CONFORMATIONS. II. A PROTON MAGNETIC RESONANCE STUDY OF THE EFFECTS OF pH ON NUCLEOSIDES AND NUCLEOTIDES IN AQUEOUS SOLUTION

It has been proposed that the average rotation conformation of the base of nucleotides and deoxynucleotides may be closely correlated with the type of ring puckering in the furanose residue, and conversely any perturbation which affects the average rotational conformation of the base may induce a conformational change in the ribose ring[1]. Since the coupling constants of ribose protons are sensitive to the changes of ribose conformation and their chemical shifts are also a good measure of ring-base interactions, any perturbation arising from solvent on the average rotational conformation can be properly monitored by the changes of these two sets of parameters.

Experimentally it was observed that both the chemical shifts and coupling constants of ribose protons in some guanine nucleosides and nucleotides change significantly with solution pD's, while others are relatively insensitive to the solution pD changes. This would suggest that the protonation of the guanine base at low pD's induces a strong ring-base interaction which will result in the changes of ribose conformation and then indirectly the changes of average rotational conformation between the base and furanose residue.

In order to understand how the protonation of different bases affects differently on the ribose conformation, we have carried out a detailed conformation analysis from ribosal proton spectra of various nucleosides and nucleotides under three different pD's. The results are shown in Table I and II.

In the case of guanine nucleosides and nucleotides, including guanosine, 1-methylguanosine, 3′-GMP and 5′-GMP, the coupling constant between $H_{1'}$ and $H_{2'}$ protons of ribose moiety, $J_{H_{1'}-H_{2'}}$, increases (Fig.1) progressively with increasing pD from 3.5 Hz at pD 2.0 to 6.2 Hz at pD 7.0. Similar changes on coupling constants and chemical shifts of other ribosal protons were also observed, especially the coupling constants $J_{H_{2'}-H_{3'}}$ and $J_{H_{3'}-H_{4'}}$, and the chemical shifts of $H_{1'}$, $H_{6'}$, $H_{5'}$ and $H_{5''}$ protons. This would indicate that protonation of this base occurring predominantly on N-3 or N-7 positions of guanine base, the former being closer to the ribose protons, would cause a strong electrostatic interaction between the base and the ribose moiety in addition to the minor influence on the ring due to the changes of electronic structure of the base upon protonation. For 5′-GMP, a significant downfield shift of $H_{5'}$ and $H_{5''}$ protons of ribose was observed at low pD's. A wider separation of thest two resonances at low pD's also indicates that the strong magnetic nonequivalence between them arises from the protonation of guanine base. It is known from Karplus relation[2],[3] that the coupling constant between the vicinal protons is related to their observed dihedral angle ϕ by

$J_{obs} = J_o \cos^2 \phi + J_1$ where J_{obs} is the observed coupling constant, J_o and J_1 are constants. The gradual shift in coupling constants of ribose protons due to base protonation may imply direct changes in dihedral angles between vicinal ribose protons or ribose conformation. The average ring base rotational conformation about C-N glycosylic bond should therefore be shifted in such a way as to enhance this ring-base interaction, making the syn-conformation more favorable at low pD's. At high pD's, the deprotonation of the base will reduce this ring-base interaction, the strain of the ribose conformation due to this electrostatic protonation is then released. The average rotational conformation about C-N glycosylic bond is more or less sterically controlled. So that anti-conformation becomes more favorable at high pD's for these compounds.

While for other nucleosides and nucleotides, such as 2-N, N-dimethyl-methylene-amino guanisine, adenosine, 5'-AMP, 2'-deoxy and pyrimidine nucleosides as well as nucleotides etc., $J_{H_{1'}-H_{2'}}$ increases only at most 1 Hz when the solution pD increases from 2.0 to 7.0. The coupling constants for the other protons on the furanose residue do not show any significant changes with pD either, compared to the case mentioned. This indicates that the ring-base interaction arising from the protonation of the base in these compounds is so weak that the ribose conformation keeps essentially unchanged for the whole pD range (anti-conformation is more favorable). For 2-N,N dimethylamino-methylene guanosine, protonation can occur only on N-7 position of the base which is relatively far away from the ribose. As a result, the effect of the

protonation does not give any significant change on the ribose conformation. The rotation about C-N glycosylic bond is also restricted due to the presence of the bulky side chain on the base, limiting the variation of the average rotational conformation to a narrow range. In the case of 2′-GMP, the steric hindrance of the 2′ phosphate group will also restrict the variation of average rotational conformation to a small range. Judging from little change in coupling constants of ribosal protons, we see that the effect of the base protonation on the ribose conformation is small in this compound. For the 2′-deoxy nucleosides and 2′-deoxy nucleotides, such as 2′ deoxy-guanosine and 5′d-GMP, the coupling constants of ribose protons are not sensitive to the protonation of the base compared to those of nucleosides and nucleotides mentioned. This would indicate that the backbone of deoxy ribose moiety is more rigid than that of ribose, and the base protonation will not give any sizeable changes on the conformation of deoxy ribose. Due to the steric hindrance arising from the rigidity of backbone of deoxyribose, the variations of the average rotational conformation about glycosylic bond upon base protonation is therefore limited.

In the case of the adenosine and 5′-AMP, the protonation of the base would be more favorable on N-1 or N-7 positions which are far away from the ribose backbone. Judging from the changes in chemical shifts and coupling constants of the ribose protons upon base protonation, the ring-base electrostatic interaction would be too weak to affect the ribose conformation, making anti-conformation more favorable for the whole pD range.

In the case of pyrimidine nucleosides and nucleotides, the protonation of the base has little effect on the ribose conformation and the rotation about glycosylic bond is more sterically hindered compared to those of purine base. As a result, anti conformation is favorable for the whole pD range.

If the conformation of ribose is a proper function of the observed coupling constants of ribose protons and if the changes in its conformation are closely correlated to the changes of the distribution of the rotational conformation C-N glycosylic bond, any perturbation on the molecule which can change this distribution will in turn change the coupling constants of ribose protons. Since the distribution of the rotational conformations about C-N glycosylic bond is Boltzmann factor or temperature controlled, the temperature study of the coupling constants of ribose protons provides additional information about this plausibility. Experimentally it was observed that the coupling constants of ribose protons are temperature dependent either at high or low pD's. The results for the temperature study of $J_{H_{1'}-H_{2'}}$ of 1-methylguanosine and 5'-GMP are depicted in Table III. If the average rotational conformation of these compounds is more favorable in syn-conformation at low pD's, the increase in temperature will tend to average the distribution of rotational conformations. This could cause the observed change in $J_{H_{1'}-H_{2'}}$ of 1-methylguanosine from 3.7 Hz to 4.8 Hz as the temperature increases from 20°C to 80°C at pD 2.2. If the anti-conformation is more favorable at high pD the syn/anti ratio will increase with increasing

temperature, in agreement with the observed decrease in $J_{H_1'-H_2'}$ from 6.2 Hz at 30°C to 5.6 Hz at 80°C for 1-methylguanosine as well as 5'-GMP in neutral aqueous solution.

From this study we have concluded that guanosine, 5'-GMP and 3'-GMP do exist in different average rotational conformation from other nucleosides and nucleotides when guanine base is protonated at low pD's.

REFERENCES

1. J. H. Prestgard and S. I. Chan, J. Am. Chem. Soc., 91,2843(1969).

2. M. Karplus, J. Am. Chem. Soc., 85,2870(1963).

3. A. Abragam, J. Chem. Soc., 3699(1962).

I. CHEMICAL SHIFT[a]

	H_8	$H_{1'}$	$H_{2'}$	$H_{3'}$	$H_{4'}$	$H_{5'}$	$H_{5''}$
1-Methyl Guanosine							
pD 7.4	4.768	2.677	1.527	1.208	1.022	0.693	0.623
4.4	4.799	2.677	1.537	1.220	1.035	0.707	0.629
2.2	5.804	2.850	1.525	1.221	1.090	0.780	0.676
2-N,N-Dimethylamino Methylene Guanosine							
pD 7.8	4.863	2.791	1.527	1.177	0.996	0.657	0.580
4.5	4.918	2.791	1.532	1.195	1.009	0.671	0.598
2.2	5.464	2.968	1.489	1.1818	1.014	0.693	0.598
2′ GMP							
pD 7.5	4.850	2.850	1.907	1.388	1.055	0.691	0.636
4.4	4.864	2.855	1.964	1.370	1.076	0.717	0.660
2.5	5.818	3.000	2.00	1.354	1.122	0.763	0.704
5′ GMP							
pD 7.5	5.036	2.722	1.554	1.300	1.136	0.850	0.830
4.4	5.045	2.772	1.527	1.277	1.162	0.942	0.915
2.5	5.827	2.890	1.503	1.295	1.223	1.088	0.977
5′ AMP							
pD 7.5	5.455	2.922	1.574	1.331	1.208	0.967	0.940
4.5	5.259	2.936	1.513	1.304	1.196	0.981	1.004
3.0	5.532	3.032	1.590	1.343	1.225	1.002	0.975
Adenosine							
pD 2.5	5.400	2.980	1.598	1.287	1.115	0.767	0.676
8.0	5.154	2.890	1.670	1.274	1.186	0.779	0.680

a :Reference; Tetramethyl ammonium chloride.

COUPLING CONSTANT (II)

	$J_{H_1', H_2'}$	$J_{H_2', H_3'}$	$J_{H_3', H_4'}$	$J_{H_4', H_5'}$	$J_{H_4', H_5''}$	$\frac{1}{2}\lvert J_{H_5'-P} + J_{H_5''-P}\rvert$	$J_{H_2'-P}$	$J_{H_5', H_5''}$
1-methyl guanosine								
pD 7.5	6.2	5.5	3.5	2.8	4.0			12.5
4.5	5.5	5.0	4.2	2.8	4.0			12.5
2.2	3.75	4.5	5.5	2.8	4.0			12.5
2-N,N-dimethyl amino methylene guanosine								
pD 7.8	5.7	5.8	4.2	3.0	4.0			12.5
4.5	5.6	5.6	4.4	3.0	4.0			12.5
2.2	4.8	5.4	4.6	3.0	4.0			12.5
2' GMP								
pD 7.5	5.4	6.0	3.8	3.0	4.0		9.5	12.5
4.4	5.4	5.6	4.0	3.0	4.0		9.0	12.5
2.5	5.2	5.4	4.4	3.0	4.0		9.0	12.5
5' GMP								
pD 7.5	6.2	5.4	4.0	2.5	3.5	4.5		12.5
4.4	5.4	5.0	4.4	2.5	3.5	4.5		12.5
2.5	3.8	4.5	5.5	2.5	2.5	5.0		12.5
5' AMP								
pD 7.5	6.0	5.3	3.7	3.0	4.0	4.0		12.0
4.5	5.5	5.2	3.6	3.0	4.0	4.0		12.0
3.0	5.4	5.0	3.5	3.0	4.0	4.0		12.0
Adenosine								
pD 2.5	5.4	5.4	4.2	3.0	4.0			12.5
8.0	6.0	6.0	3.5	3.0	4.0			12.5

TEMPERATURE STUDY OF $J_{H_1'-H_2'}$ (III)

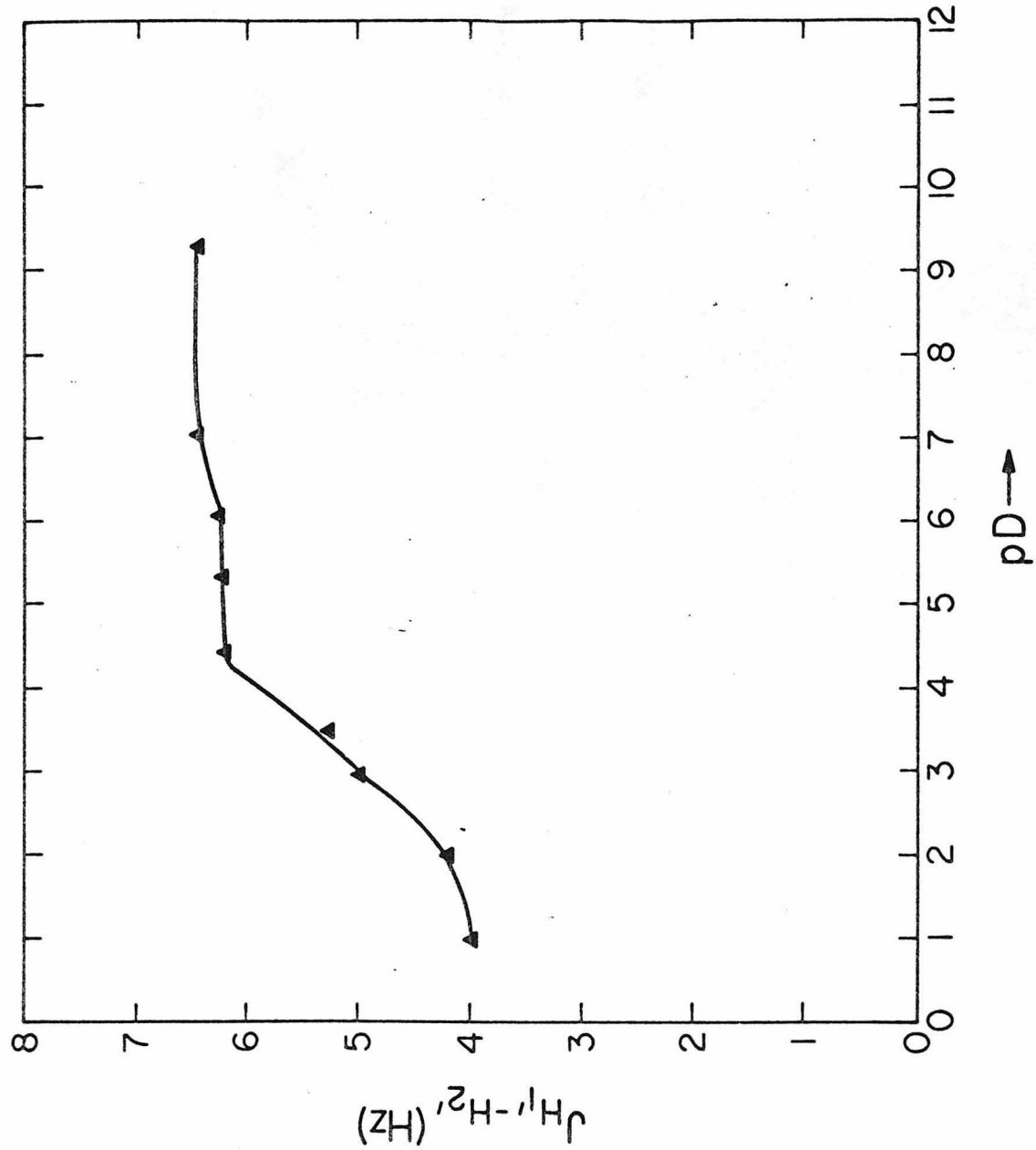
1-methyl guanosine

pD	Temperature	$J_{H_1', H_2'}$
2.2	30°C	3.8 Hz
2.2	60°C	4.6 Hz
2.2	80°C	4.8 Hz
4.5	30°C	5.5 Hz
4.5	80°C	5.6 Hz
7.4	30°C	6.2 Hz
7.4	80°C	5.8 Hz

5′ GMP

1.5	18°C	3.4 ~ 3.5 Hz
1.5	60°C	4.1 ~ 4.2 Hz
7.8	18°C	6.2 Hz
7.8	60°C	5.6 Hz

Fig.1. PD dependence of $J_{H_1'-H_2'}$ of guanosine


$J_{H_1'-H_2'}$ (Hz)
pD
0
1
2
3
4
5
6
7
8
9
10
11
12