

## STUDIES RELATED TO PRIMITIVE CHEMISTRY

(A preliminary investigation of the applicability of proton and nitrogen-14 nuclear magnetic resonance to the characterization of interactions between amino acid and nucleic acid constituents and a discussion of previous work on these interactions and their possible relation to prebiotic synthesis of biopolymers.)

by

S. L. Manatt, E. A. Cohen, A. M. Shiller and S. I. Chan  
Department of Chemistry and Chemical Engineering  
California Institute of Technology  
Pasadena, California 91109  
December 28, 1973

NASA-CR-138033)	STUDIES RELATED TO	N74-19771
PRIMITIVE CHEMISTRY.	A PRELIMINARY	
INVESTIGATION OF THE APPLICABILITY OF		
PROTON AND NITROGEN-14 NUCLEAR (California		Unclas
Inst. of Tech.)	100 p HC \$8.00 CSCL 07C	G3/06 16618

FINAL REPORT

Grant NGR 05-002-314 through Ames Research Center.

TABLE OF CONTENTS

ABSTRACT . . . . .	ii
INTRODUCTION . . . . .	1
SUMMARY OF PREVIOUS WORK . . . . .	5
SCOPE OF PRESENT WORK . . . . .	21
RESULTS AND DISCUSSION . . . . .	27
Proton NMR Studies . . . . .	27
<sup>14</sup> N NMR Studies of Several Amino Acids . . . . .	75
EXPERIMENTAL . . . . .	84
CONCLUSION . . . . .	95

STUDIES RELATED TO PRIMITIVE CHEMISTRY

ABSTRACT

Preliminary proton nuclear magnetic resonance (NMR) studies have been made to determine the applicability of this technique for the study of interactions between monomeric and polymeric amino acids with monomeric nucleic acid bases and nucleotides. Proton NMR results for aqueous solutions ( $D_2O$ ) demonstrated interactions between the bases cytosine and adenine and acidic and aromatic amino acids. Solutions of 5'-AMP admixed with amino acids exhibited more complex behavior but stacking between aromatic rings and destacking at high amino acid concentration was evident. The multisite nature of 5'-AMP was pointed out. Chemical shift changes for adenine and 5'-AMP with three water soluble polypeptides demonstrated that significant interactions exist. The assignment of the  $H_3$  proton as the upfield one in adenine has been made for the first time. A discussion of 1:1 association and self-association is given. A spectrometer for recording  $^{14}N$  NMR spectra was assembled. A discussion on how  $^{14}N$  NMR linewidths of amino acids can give information on weak association is given. A study of the linewidth versus pH dependence of several amino acids is reported for the first time. It was found that the linewidth-pH profile of each amino acid is unique.

It is concluded that NMR techniques can give significant and quantitative data on the association of amino acid and nucleic acid constituents. A review of the literature on these interactions is given. A number of possible future experiments are suggested for further characterization of such interactions.

### INTRODUCTION

Research from a number of laboratories has demonstrated that model primitive atmospheres and surface conditions can produce all the classes of simple molecules which make up the biopolymers of living organisms.<sup>1</sup>

---

1. For review and recent work see: (a) A. L. Miller and H. C. Urey, Science, 130, 245 (1959); (b) H. H. Pattee, Adv. Enzym., 27, 381 (1965); (c) C. Ponnamperna, Quart. Rev. Biophys., 4, 2 & 3, 77 (1961); (d) J. S. Hubbard, J. P. Hardy and N. H. Horowitz, Proc. Nat. Acad. Sci. U.S., 68, 574 (1971).

---

Although a number of reasonably simple chemical and physical mechanistic schemes have been proposed to explain the observed syntheses of simple organic molecules from inorganic starting materials, the details of the subsequent accumulation and modification of the initially formed molecules to yield polymers possessing well-defined sequences and possessing catalytic and replicating properties are not known. Implicit in many arguments regarding whatever mechanisms determined the chemical evolution that took place is that from the tremendous number of classes of organic and inorganic compounds, which can be produced by the incidence of various types of energy on the primitive earth, a certain set of classes and certain sets of isomers within classes or types of compounds appear to possess qualities and relationships which made them predominate as the building blocks for living cells.

For some time it has been suggested that there should be some definite stereochemical relationship between the various amino acids and the base sequence of their codons and/or anticodons (the so-called "stereochemical theory").<sup>2-6</sup> These relationships are believed to have been important in

- 
2. L. Pauling and M. Delbruck, Science, 92, 77 (1940).
  3. G. Gamow, Nature, 173, 318 (1954).
  4. (a) C. R. Woese, Proc. Nat. Acad. Sci. U.S., 59, 110 (1958);  
(b) C. R. Woese, "The Genetic Code," Harper and Row, New York, 1967;  
(c) C. R. Woese, J. Mol. Biol., 43, 235 (1969).
  5. F. H. C. Crick, J. Mol. Biol., 38, 367 (1968).
  6. L. E. Orgel, J. Mol. Biol., 38, 381 (1968).
- 

the evolution of the present day genetic and protein synthesis apparatus. However, as yet no definitive experiments elucidating and clarifying these relationships in terms of reasonable schemes of chemical evolution of the translation apparatus have been accomplished. Woese<sup>2a</sup> has pointed out that it is possible that the nucleic acid directed protein synthesis as it is understood today is so many evolutionary steps removed from its origin that the two have little or nothing in common, so that present knowledge could strongly prejudice experimental and theoretical approaches to understanding this evolution. In spite of this possibility, it has been strongly argued that it is essential to pursue the stereochemical theory and acknowledged that development of a fruitful approach may require a significant amount of initial probing.<sup>4-6</sup>

In this report we first present a discussion of much of the published work regarding amino acid and nucleic acid constituent interactions. Most of these works concern objectives directed towards elucidation of current day biochemical mechanisms and involve interactions between pairs of biopolymers or a polymer and a small molecule.

Besides the potential significance of amino acid-nucleic acid interactions in biochemical evolution, they are of prime importance currently relative to general enzyme-nucleic acid interactions that determine the mechanisms of action of histones, repressors, initiators, polymerases, various drugs and the detailed understanding of the role of t-RNA's in translation and protein synthesis.

It is our conviction that the detailed physical chemical understanding of the possible interactions between these classes of biomolecules at the monomer and small oligomer levels must precede a successful understanding at the more complicated level. The experimental work which will be discussed in this report is directed mainly at detection of interactions at the monomer level using the NMR technique. Throughout this report, suggestions are made regarding subsequent experiment bearing on the fuller characterization of these biomolecule interactions. The present work was envisaged as a preliminary probing study which was directed towards establishing the scope and feasibility of NMR techniques in searching for evidence relating to the chemical evolutionary relationships and interactions between amino acid and nucleic acid constituents. Although the biopolymers initially responsible for the first replicating cells could

be considerably different in certain aspects than what we recognize today, it is generally agreed that monomeric amino acids, oligopeptides, monomeric nucleic acid bases, nucleosides, nucleotides, oligonucleotides and the same physical chemical principles existed on the primitive earth as do today, so that NMR experiments should give relevant information.

SUMMARY OF PREVIOUS WORK

The first work concerned with trying to characterize nucleic acid-protein interactions investigated the precipitates formed between DNA and histones or polylysine.<sup>7,8</sup> All these studies were very qualitative and

- 
7. E. Chargoff, C. F. Crampton and R. L. Lipshitz, Nature, 172, 289 (1953);  
C. F. Crampton, R. Lipshitz and E. Chargoff, J. Biol. Chem., 211, 125 (1954);  
P. Spitzik and E. Chargoff, ibid., 215, 765 (1955).  
8. J. A. Lucy and J. A. V. Butler, Biochim. Biophys. Acta, 16, 431 (1955).

---

demonstrated formation of complexes which could be redissolved in 1-2 M NaCl solution. From chromatographic experiments with the solubilized complexes it was shown that G-C rich DNA elutes faster than A-T rich DNA which was presumed to be more extensively complexed with histone or polylysine.

Subsequently, investigation of interactions of RNA with polylysines and polyarginine, however, showed that these peptides had higher affinity for G-C rich region of RNA in contrast to the A-T preference with DNA.<sup>9</sup> Most

- 
9. H. A. Sober, S. Schlossman, A. Yaron, S. A. Latt, and G. W. Rushizky, Biochemistry, 5, 3608 (1966).

---

studies still concentrated on the interactions of DNA with polypeptides and more detailed information began to emerge. Thus, it was shown that the interactions with DNA were cooperative and reversible under most circumstances.<sup>10-13</sup> However, non-uniform binding and irreversibility were suggested



- 
10. M. Leng and G. Felsenfeld, Proc. Nat. Acad. Sci. U.S., 53, 1325 (1966).
  11. E. D. Akinrimisi, J. Bonner and P. Ts'o, J. Mol. Biol., 11, 128 (1965).
  12. O. E. Olins, A. L. Olins, and P. H. von Hippel, J. Mol. Biol., 24, 157 (1967).
  13. M. Tsuboi, K. Matsuo, and P. O. P. Ts'o, ibid., 15, 251 (1966).
- 

with polylysine under conditions of large local excess of polylysine. Sober in careful equilibrium dialysis studies with RNA and individual oligomers of (L-lysine)<sub>n</sub>- $\epsilon$ -N-DNP-L-lysine (n = 3,4,5,6,7 and 8) showed that there was a 1:1-lysine:phosphorus ratio in the soluble complexes, that binding was stronger to poly(IC) than to poly(AU) and that the total binding energy and difference between binding energies increased linearly with oligolysine chain length.<sup>14</sup> However, even more important Sober demonstrated

- 
14. S. A. Latt and H. A. Sober, Biochemistry, 6, 3293 (1967).
- 

the importance of cations in affecting the extent and specificity of binding of oligolysines and RNA and suggested that water of hydration might be very important in control of nucleic acid-protein interactions.<sup>14,15</sup> Presumably,

- 
15. S. A. Latt and A. A. Sober, ibid., 6, 3307 (1967)
- 

a similar situation applies to DNA.

Polylysine-DNA complexes were studied again by Felsenfeld in sedimentation velocity, light-scattering, optical rotatory dispersion and circular dichroism experiments.<sup>16</sup> From these results it was concluded that the complexes

---

16. J. T. Shapiro, M. Leng and G. Felsenfeld, ibid., 8, 3219 (1969).

---

were highly solvated and "remarkably" uniform in size with average radius of about 1700 Å and that considerable perturbation of DNA secondary structure exists or the complexes involve formation of structure with long-range order. From all the work on these complexes which had been done to this point the following general conclusions were drawn:

1. DNA's and RNA's, both mammalian and bacterial, form soluble and/or insoluble complexes with basic polypeptides.
2. Such complexation increases the stability of nucleic acids to thermal denaturation and the stability varies with the nature of the polypeptide.
3. Interaction with DNA is cooperative and that with RNA probably is too.
4. Complexes appear to have definite stoichiometry; i.e., amino acid residue: base ratio of 1:1.
5. Complex formation produces significant spectral changes.

A number of more recent studies have been made on complexes of polylysine and polyarginine with DNA but little new specific structural information has emerged. It has been found that anomalous circular dichroism (CD) spectral changes occurring for DNA-polylysines are similar to those observed for DNA-lysine-rich histones. In the case of polyarginine, however, much less altered CD spectra are observed. These later changes are similar to what is observed with arginine-rich histones and protamine.<sup>17</sup> Polylysine complexes

---

17. D. Carroll, Biochemistry, 11, 421 (1972) and reference cited therein.

---

with poly(U) and converts the polynucleotide to its double-stranded form. However, polyarginine and protamine are ineffective, as indicated by CD, in complexing with poly (U).<sup>18</sup>

---

18. D. Carroll, Biochemistry, 11, 426 (1972).

---

Several types of data which are still missing are any structural studies of these complexes or data on dynamical aspects of these complexes such as correlation times and substrate exchange rates.<sup>18a</sup>

---

18a. For interactions of histones with nucleic acid constituents, the rate appears slow; S. I. Chan, unpublished work.

---

Parallel with studies of polypeptide-nucleic acid interactions, both interactions between mononucleotides and polypeptides and amino acid derivatives and polynucleotides have been sought. Perhaps the simplest approach was Woese's experiments in which he sought interactions between amino acids and various organic bases via their effect on paper chromatographic retention times.<sup>19-21</sup> For example, from plots of log amino acid

---

19. C. R. Woese, D. H. Dugre, W. C. Saxinger and S. A. Dugre, Proc. Nat. Acad. Sci. U.S., 55, 966 (1966).

20. Ref. 4b, pp. 150-178.

21. C. Woese, D. Dugre, S. Dugre, M. Kondo and C. Saxinger, Cold Spr. Harb. Symp. Quart. Biol., 31, 723 (1966).

---

retention times vs. log mole fraction water for series of pyridine-methyl substituted pyridine: water mixtures he argued that mainly polar interactions were present, but additionally, some non-polar interactions were also present. He concluded that the polar interactions involved the ring

nitrogen atoms interacting with the polar portions of the amino acids "plus a non-polar interaction, between the more 'organic' portions of the bases and amino acids". Although his experimental evidence is not compelling, he has postulated some general features regarding the genetic code and these interactions which have stimulated much debate:<sup>5,6</sup>

1. The general geometry of any codon (or anticodon)-amino acid complex is similar for all possible interaction pairs.
2. There exists a hierarchy of position in a codon or anticodon, defined both in terms of kind of interaction (polar vs nonpolar) and strength of interaction.
3. Total interaction between an amino acid and codon or anticodon for all pairs is about the same but partition between polar and nonpolar interaction differs.
4. It is conceivable that translation evolved out of a different sort of nucleic acid-amino acid or peptide relationship than we know today so that our knowledge of modern translation may seriously blind search for understanding of the evolution of the genetic code.

The most interesting results bearing on these later features seems to be the ordering of  $\underline{t}$ -RNA's from a particular organism by various separation procedures. Woese<sup>21</sup> found that the order of their elution from columns of methylated albumin on kieselgahr and their counter current distributions follows some definite patterns. For the  $\underline{t}$ -RNA's related to an RNA codon having a common base in the 2nd codon position, it was observed with few exceptions that those with  $G_I$  codon elute first, then those with  $A_I$

then those with C<sub>I</sub> and finally those with U<sub>I</sub>. Similar elution patterns have also been observed for Sephadex systems.<sup>22</sup> The implication of these

---

22. S. G. Nathenson, F. C. Dohan H. H. Richards and G. L. Canteni, Biochemistry, 4, 2412 (1965).

---

results is that t-RNAs' stereochemical properties fall into classes related to amino acids and the 2nd codon.<sup>21</sup> It would seem that these observations should be followed up.

Wagner and Arav<sup>23</sup> examined the interaction of various mononucleotides

---

23. K. G. Wagner and R. Arav, Biochemistry, 7, 1771 (1968).

---

with poly-L-lysine and poly-L-arginine by equilibrium dialysis. They concluded that at neutral pH the observed interactions were mainly non-specific electrostatic ones involving the negatively charged phosphate groups of the nucleotides and the positively charged side chains of the polypeptides. However, they also believed that there was evidence for a component of a more specific non-electrostatic effect depending on the nature of the nucleotide base. This conclusion was based on an approximate correspondence between the Michaelis constant of 3'-ribonucleoside monophosphates with a certain 3'-nucleotidase<sup>24</sup> and the apparent association

---

24. H. S. Loring, J. E. McLennon and T. L. Walters, J. Biol. Chem. 241, 2876 (1966).

---

constants of the 5'-ribonucleoside monophosphates with poly-L-arginine which they studied. They believed that, because the direct interaction between a nucleotide and its binding site is involved in substrate recognition of enzymes dealing with nucleotides and because 3'-nucleotides in this case are recognized regardless of the base attached, the contribution to the Michaelis constant represents relatively unspecific affinities which appeared to be comparable to the factors giving rise to the binding behavior changes they observed.

Light scattering photometry has been used to study the interaction between various poly-L-lysines and the nucleotides and some related derivatives by Lacey.<sup>24a</sup> These experiments at low ionic strength, so as to

---

24a. J. C. Lacey and K. M. Pruitt, Nature, 223, 799 (1969).

---

maximize ionic interactions, showed that extreme turbidity was produced on addition of mononucleotides except 5'-UMP, to polylysine, mol. wt. 100,000. However, addition of the corresponding nucleosides did not produce turbidity. In the case of polylysine, mol. wt. 3,000, however, only 5'GMP produced turbidity. Careful microscopic examination of the turbid solutions revealed that the turbidity arises from tiny spheres which could be centrifuged out of solution. These results were interpreted to show that the effectiveness in producing turbidity is in the same order as the tendency of the mononucleotide to self-associate. A two step process producing the turbidity was suggested: a) ionic binding between the nucleotide and polylysine followed by b) interchain nucleotide-nucleotide interactions of polypeptide bound nucleotides. This scheme

was supported by ionic strength, temperature, and stoichiometry studies at various pH's. Although these results were felt to reveal no direct relationship between peptide sequence and codon specificity, several more appropriate studies were suggested. Thus, studies of mononucleotide-polypeptide mixtures using amino acid copolymers with fewer charges might be more revealing. In conclusion, Lacey strongly suggested a model for the origin of the genetic code which involves initial formation of polypeptides followed by specific complexing between the latter and nucleotide constituents. Held in the complex, polymerization takes place somehow to produce polynucleotide in the complex. These complexes dissociate at higher temperatures and the polyamino acid proceeds to act as a template for further polyribonucleotide synthesis. Leading to this conclusion, Lacey built models which indicated a reasonable scheme by which it was possible to get the genetic code stoichiometry of three mononucleotides to one amino acid residue in a polymer. In this model for each codon, the amino acid side chain lies between the first and second nucleotides of the codon. The third nucleotide is remote from the amino acid side chain and might be expected to possess little interaction with the amino acid residue. This situation is consistent with many other observations on the present day genetic code which suggest the third nucleotide in a base triplet is relatively unimportant for codon specificity.<sup>4,5,21</sup> The

suggestion that polypeptides were first formed in significant amounts in the prebiotic environment is based on much evidence that polypeptides are much more readily formed from amino acid monomers under simulated prebiotic conditions than polynucleotide from mononucleotides.<sup>4c,5,6,29,30</sup>

Some ORD studies have been made on the interaction between polylysine, polyornithine and polyarginine with the nucleotide monophosphates in water and 60% dioxane.<sup>25</sup> It was concluded that the nucleotides can

---

25. J. M. Rifkind and G. L. Eichhorn, Biochemistry, 9, 1753 (1970).

---

stabilize the helical conformation of peptides to a degree which corresponds to the base-stacking ability of the nucleotides. Also, it was suggested that the nucleotides differ in the extent by which they undergo conformational changes when oriented around polypeptides. It was pointed out that the geometry for base stacking around a helical polypeptide is rather poor. The data were interpreted to indicate that GMP is the only mononucleotide which complexes and maintains significant stacking interactions. Although spectral changes were observed with CMP and AMP, they were not appropriate to stacking interaction. The helical conformation of polyarginine is stabilized more readily than that of polylysine but the stacking interactions of GMP occur less readily around polyarginine than around polylysine. It was suggested that polyarginine's guanidinium group prefers to interact



with two of the oxygens of the phosphate in such a manner that the base stacking is more difficult.

Several authors have sought interactions between aromatic amino acids and nucleic acid components. For example, Helene studied frozen aqueous solutions of tryptophan with nucleosides by reflectance and luminescence spectroscopy.<sup>26</sup> It was suggested that really the cell is not like a dilute

---

26. T. Montenay-Garestier and C. Helene, Biochemistry, 10, 300 (1971).

---

aqueous solution and that an enzyme-nucleic acid interaction may be mainly strong specific interactions, while the transportation to the active site involves mainly strong nonspecific electrostatic forces. Thus close contact between molecules can be promoted by freezing which allows aggregates of solute molecules to form in cavities. Helene found new absorption at longer wavelengths and new fluorescence spectra characteristic of weak charge-transfer or electron donor-acceptor complexes which would be consistent with an insertion of indole ring of the tryptophan in a stacking scheme involving the purine or pyrimidine ring of the nucleosides.

Ponnamperuma has studied the interactions of nine column-bound amino acids with ribonucleoside 5'-phosphates in equilibrium studies.<sup>27</sup>

---

27. W. C. Saxinger and C. Ponnamperuma, J. Mol. Biol., 1, 63 (1971).

---

It was suggested that the amino acid bound to a cross-linked polyvinyl amine is analogous to an N-terminal protein amino acid and possibly to

biologically activated amino acids. Although there were a number of features of these studies which were unclear, it was concluded that nucleotide binding to amino acids (even the non-aromatic ones) follows roughly the tendency of bases to stack, but a component related to the polarity of the amino acid sidechain plays a role too. A site-binding model for the nucleotide-amino acid interaction which involved base and side chain overlap was suggested. However, there seems to be some additional data needed for the interpretation of this work before its significance can be determined. The results of experiments where the various nucleotides are equilibrated with a binding system not containing any column bound amino acids needs to be known along with results with plain sugar phosphates, nucleosides and the bases themselves.

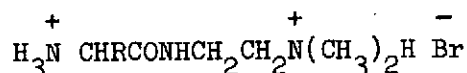
In work to characterize the difference between the interactions of L and D enantiomers with polynucleic acids, Gabbay<sup>28</sup> studied the

---

28. E. Gabbay and R. Kleinman, J. Am. Chem. Soc., 89, 7123 (1967).

---

interaction of amino acid amides of the type:



with nucleic acid. Measurable differences in  $T_m$  for D and L cases were noted, which were interpreted to mean that L stabilizes most. Also, it appeared that the degree of stabilization depends on the amino acid. Thus, increasing the size of the hydrophobic substituent R- lowered the degree of stabilization while a polar R- substantially stabilized the helix to melting.

There have been some wet chemical studies, notably those of Fox, which has been directed towards detection and characterization of possible relations between amino acids and nucleotide constituents.<sup>29,30</sup> Fox, et al.,

---

29. S. W. Fox, J. C. Lacey and T. Nakashima, in "Nucleic Acid-Protein Interactions", ed. by D. W. Ribbons and J. F. Woessner, North-Holland Publishing Co., Amsterdam, 1971, p. 133.

30. T. Nakashima and S. W. Fox, Proc. Nat. Acad. Sci. U.S.A., 69, 106 (1972). and papers cited therein.

---

have studied the interaction of the thermally formed polypeptides called proteinoids with polynucleotides in aqueous solution and found that micro-particles are formed.<sup>30</sup> It was found that particle formation required the proteinoid have a net positive charge. The facts that these particles dissociate at high ionic strength and high pH suggest that electrostatic interaction are quite important. Evidence for the possibility that the degree of interaction might relate to the major amino acid constituent(s) of a particular proteinoid in a codon or anticodonic manner for any particular homopolynucleotide was not found. Fox and co-workers have also treated the proteinoid-polynucleotide micro-particles with <sup>14</sup>C-AMP-anhydrides of several amino acids.<sup>29</sup> The extent of an amino acid's incorporation was in most cases found to relate to its present-day codon present in the homopolynucleotide in the micro-particle. However, in other experiments they observed anti-condonic relations. This has led them to suggest perhaps the recognition process might be sensitive to conditions and might be switchable

between condonic and anticondonic. The proteinoid and micro-particle system are apparently very complicated chemical mixtures which could be characteristic of the early evolutionary chemical soups which existed on earth but it appears that these systems need to be understood much more completely before significant conclusions on specific structural aspects of polypeptide-polynucleotide interactions or relationships will come from them. Perhaps, careful degradative studies of these materials might be a fruitful and informative approach.

As is discussed in more detail below, nuclear magnetic resonance (NMR) is a very powerful technique for studying weak interaction. However, there are few studies thus far on the question of the nature of the interactions of amino acids and peptides with nucleotide constituents and polynucleotides. The most detailed published NMR studies to date have been on the interactions between aromatic amino acids and their derivatives with purine and pyrimidine derivatives.<sup>31,32,33</sup> The principal interactions

- 
31. H. Claude, J. Dimicoli and F. Brun, Biochemistry, 10, 3802 (1971).
  32. J. Dimicoli and C. Claude, J. Am. Chem. Soc., 95, 1036 (1973).
  33. K. G. Wagnet and R. Lawaczeck, to be published; private communication.
- 

appear to involve stacking and destacking of the aromatic rings. It was found that the degree of interaction between various aromatic amino acid constituents and purines was the same order as that for purine self-association while it was greater than the tendency for pyrimidine self-association.

The association of various 5'-nucleotides with tryptamine and 5'-methyltryptamine has been studied by NMR.<sup>33</sup> It was found that the purine derivatives again interact more strongly than the pyrimidines with aromatic amino acid related molecules. The observed interactions were attributed to base stacking, self-association and an electrostatic contribution.

In a study of a number of amino acids esters with nucleotides and polynucleotides, increased shielding was uniformly noted in the case of mixtures with aromatic amino acid.<sup>34</sup>

---

34. C. Saxinger and C. Ponnampereuma, paper presented Fourth International Symposium on the Origin of Life, Barcelona, Spain, June 24-28, 1973; private communication from C. Ponnampereuma.

---

Poly (A) has been studied with a number of amino acids and their derivatives by NMR.<sup>31,35</sup>

---

35. M. Taszka and M. Mandel, Proc. Nat. Acad. Sci. U.S., 68, 1190 (1971).

---

It was observed that the aromatic amino acids interact with single-stranded polyadenylic acid as indicated by chemical shift changes of the adenine H<sub>2</sub> and H<sub>8</sub> protons. The changes observed were consistent with the destacking of the partly-stacked polynucleotide chain by intercalation of the amino acid aromatic rings. It was claimed that neither aliphatic amino acids nor lysine affected the H<sub>2</sub>-H<sub>8</sub> line positions or the T<sub>m</sub>, melting temperature, of the poly (A).<sup>35</sup>

In one study with a t-RNA, significant line broadening was observed in the spectra of gly-gly, gly-DL-ser and gly-DL-leu.<sup>36</sup> The observed NMR

---

36. J. M. Backer, Yu. N. Molin and A. S. Sorokin, Biochem. and Biophys. Research Comm., 38, 871 (1970).

---

line broadening was attributed to interaction between the amino groups and the negatively charged phosphate groups, as no broadening was observed for N-substituted dipeptides.

Gabbay and coworkers have studied the proton NMR spectra of over 50 peptides<sup>37</sup> in the presence of various nucleic acids. They found in

---

37. E. J. Gabbay, K. Sanford and C. S. Baxter, Biochemistry, 11, 3429 (1972).

---

general with dipeptides and dipeptide amides containing an aromatic amino acid that the aromatic residues experience upfield chemical shifts and line broadening. In addition, they noted that dipeptide amides with aromatic rings cause dramatic decrease in the specific viscosity of DNA. Gabbay and coworkers have also studied the interaction of various di- and tripeptides containing aromatic amino acids with nucleic acids of different A-T-G-C compositions.<sup>38</sup> The results suggest a definite A-T selectivity for

---

38. E. J. Gabbay, Karl Sanford, C. S. Baxter and L. Kapicak, Biochemistry, 12, 4021 (1973).

---

the aromatic amino acids: tryptophan > phenylalanine > tyrosine > leucine. It was suggested that this A-T selectivity for aromatic amino acids could have significant biological significance. The selective intercalation on DNA could be one of the mechanisms whereby polypeptide associates with polynucleotide.

In summary, a number of workers are attacking the problem of the interactions of amino acids and polypeptides with nucleic acid and their bases. Only recently have specific, detailed information begun to be obtained. However, still many details are lacking on how what is known can be related to the current mechanisms of gene directed protein synthesis and the evolutionary mechanisms for the accumulation of abiotically synthesized compounds. The importance for seeking some specific stereochemical relationships between protein and nucleic acids is still acknowledged time and time again as an important problem, but the lack of direct, concrete experimental evidence still exists. It is recognized that experiments along these lines should be most carefully designed and complete than previous ones to avoid the pitfall of ambiguity.<sup>4,5,6,30</sup>

SCOPE OF PRESENT WORK

Several types of interactions between amino acid derivatives and polymers and nucleic acid constituents are conceivable. For example, the phosphate groups of nucleic acids are certainly involved in significant electrostatic interaction between basic amino acid side chains. It is well known that hydrogen bond formation between purine and pyrimidine bases is an extremely important factor in the determination of the DNA and RNA conformation.<sup>4b,39,40</sup> There is every reason to expect that this type of

---

39. J. D. Watson and F. H. C. Crick, Nature, 171, 737 (1953).

40. J. D. Watson, "Molecular Biology of the Gene", W. A. Benjamin, New York (1965).

---

interaction might exist between amino acid derivatives and nucleic acids derivatives. It is conceivable that hydrogen bonding could exist not only between amino acid component and the nucleic acid bases, but also to the sugar and phosphate groups of the nucleic acid. There is much previous work from ours and other laboratories which has demonstrated that base stacking interactions involving aromatic rings of nucleic acid constituents is, in addition to base pairing, a very important factor in determining polynucleic acid structure and properties.<sup>41</sup> The nature and extent of any

---

41. See, for example, G. P. Kreishman and S. I. Chan, Biopolymers, 10, 159 (1971) and ref. cited therein.

---



similar interaction involving aromatic amino acids and nucleic acid has only recently been recognized as discussed above. Finally, it has been suggested that interaction involving non-polar amino acid side chains with nucleic acid constituents may exist and could have been important in evolution.<sup>19,21,27,34</sup> However, this type of interaction we regard as being highly improbable from the simple fact that nucleic acids (even alkyl substituted ones which are minor components of RNA's and DNA's) are highly polar molecules with no attached hydrocarbon chains of any length such as exist in many of the amino acids. Thus, initially, we have directed our attention and consideration to searching for interactions which might arise from: a) electrostatic interactions related to the phosphate group; b) hydrogen bond formation; and c) intercalations between aromatic groups. The work reported here is some preliminary studies directed towards detection and characterization of these types of interactions between these two classes of biomolecules.

Nuclear magnetic resonance (NMR) spectroscopy has been found to be an extremely useful tool for the investigation of weak interactions. This technique possesses extreme sensitivity for detection of slight conformation and environmental changes within and between molecules. For example, proton NMR has been shown to be ideally suited for studying the intercalation interactions of nucleic acid derivatives.<sup>41</sup> Thus, from our understanding of the details of the monomeric base, mononucleoside, mononucleotide, and oligonucleotide interactions we felt that we would be in a good position to identify what prevails in the more complicated experimental systems containing an added amino acid component. The plan of attack originally

was to carry out detailed proton NMR studies of most of the amino acids monomers with added monomeric nucleic acid base, added mononucleoside, added nucleotide and then progress to some selected oligonucleotides. Similar studies were contemplated with some selected polypeptides with various nucleic acid components. Some recent results of NMR studies of nucleic possessing quadrupole moments<sup>42,43</sup> suggested that another NMR

---

42. T. R. Stengle and J. D. Baldeschwieler, Proc. Nat. Acad. Sci. U.S., 55, 1020 (1966).

43. R. G. Bryant, J. Am. Chem. Soc., 91, 976 (1969) and refs. cited therein.

---

approach to detection and characterization of weak interactions between amino acid monomers and derivatives and nucleic acid constituents might be possible by <sup>14</sup>N NMR studies of amino acids and small peptides admixed with nucleic acid constituents. Very little <sup>14</sup>N NMR has been carried out on biologically important molecules.<sup>44</sup> To explore the potential of this NMR

---

44. Professor Rex Richards, private communication, has informed us that his group at Oxford has constructed a superconducting, pulse spectrometer for <sup>14</sup>N NMR with which they have recently investigated a number of biomolecules, especially a number of the amino acids.

---

technique, some attention was directed towards demonstration of a <sup>14</sup>N NMR capability and execution of some <sup>14</sup>N NMR studies of several amino acids.

Based on the results obtained in the studies outlined above, additional studies with model compounds,  $^{31}\text{P}$  NMR studies,  $^{13}\text{C}$  studies and relaxation time measurements were anticipated during the tenure of this project. Originally it had been expected that this research contract would be composed of two parts covering a period of about eighteen months with resources for the first six months of the order of 17K and for the last year of the order of 65K. The plan discussed above was conceived accordingly. However, for many reasons only the first increment of funding and time were allotted to this research undertaking so that what was accomplished was much more limited than the plan discussed above.

Under the restraints of funding available, the following specific studies were carried out:

1. Spectra of twenty common amino acids in  $\text{D}_2\text{O}$  in presence of cytosine. These studies were carried out because, surprisingly, there are few sets of detailed spectral data for all the amino acid under comparable conditions. Thus, these results would serve to introduce us to the problems related to recording such spectra and also indicate the extent and type of spectral changes one needs to be aware of with these classes of molecules.

2. The pD dependence of cytosine spectrum and its spectra admixed with several amino acids at pD 7. This type of study is important because it is well known that the charge on aromatic rings affects their ability to stack and other possible interactions as well. The resource limitation determined that our efforts in this direction with cytosine be very limited.

3. A study of the spectrum of adenine at pD 7 in presence of five different added amino acids (one from each of several classes of amino acids, i.e., aliphatic, basic, acidic, and aromatic).

4. Spectra of 5'-AMP mixed with the same five amino acids studied in 3 above.

5. Study of adenine spectrum at pD 7 in presence of three water soluble polypeptides.

6. Spectra of 5'-AMP at pD 7 in presence of same polypeptides studied in 5 above.

7. Assembly and testing of an NMR spectrometer system for recording  $^{14}\text{N}$  spectra of liquid samples at 4.3 MHz.

8. NMR studies of the  $^{14}\text{N}$  spectra of several amino acids as a function of pH. Such a study has never been previously reported. This study is necessary as a forerunner to detection and characterization of weak interaction involving amino acids by  $^{14}\text{N}$  NMR.

There were several situations which we felt might arise in the proton NMR studies contemplated above. There would be the case of complex formation involving either two small molecule components or a small molecule and an oligomer or high polymer which would only involve chemical shift changes of both components. This situation would be characteristic of fast exchange and rapid tumbling of the complex on the NMR time scale. A second situation would give rise to broadened signals and shifts of one or both components. If both components give rise to observable signals, then this situation would indicate slow exchange and/or reduced molecular tumbling in the complex. A third situation might be if the exchange is slow but the complex tumbles

fast on an NMR time scale so that each of the components and the complex give separate signals. A fourth situation would be if one component exhibits very broad lines or no visible signal and the small molecule component gives sharp but shifted signals. This situation would indicate fast exchange and fairly fast tumbling and is very similar to the first situation above. A fifth situation would involve one component exhibiting very broad or no visible signals and the small molecule component showing broadened signals and shielding changes. This situation would be characteristic of some restriction of molecular tumbling of the complex (and thus the polymer) and/or slow exchange, both approaching the NMR time scale. Finally, for a rigid macromolecule-small molecule complex where the small molecules experiences very restricted motion, a very broad signal would be observed. This would be true for both slow and fast exchange. In the case of the  $^{14}\text{N}$  studies of the amino acids, we expected that the change of structure of the amino acids with pH would modulate the electric field gradients at the nitrogen nuclei of the amino acids. These changes, along with changes of exchange rate, would change the nitrogen nuclear quadrupole relaxation which determines the  $^{14}\text{N}$  NMR linewidth. Thus, substantial changes of line width were expected with changes of pH. In the case of mixtures of an amino acid and a nucleotide constituent, one might expect to see line width changes indicative of modulation of the  $^{14}\text{N}$  line width by weak complex formation.

With these possible situations in mind, we approached the proton spectral work and  $^{14}\text{N}$  NMR spectral work suggested above.

## RESULTS AND DISCUSSION

### Proton NMR Studies

The first proton NMR studies in this work recorded the spectra of cytosine in the presence of various amino acids. The spectrum of cytosine in  $D_2O$  consists of a pair of doublets whose centers are separated by about 152.86 Hz at pD 7.0 and 0.0279 M. Their chemical shifts are down field from the tetramethylammonium (TMA) ion internal standard by 430.90 Hz for  $H_6$  and 278.04 Hz for  $H_5$ . In the initial experiments, a cytosine solution of known concentration was added to weighed quantities of various amino acids. The quantity of amino acid was that approximately equivalent to its water solubility for the least soluble ones and about 1 M for the soluble ones. Table I summarizes the spectral data from these mixtures and Figs. 1 and 2 show the chemical shift data plotted against concentration. This data shows that there are only small changes of the cytosine  $H_5$  and  $H_6$  chemical shifts with most amino acids and as a function of amino acid concentration except in the case of the acidic amino acids, glutamic acid and aspartic acid, which show significant down field shifts, and for the aromatic acids, phenylalanine and tryptophan, which show upfield shifts.

To firmly establish the origin of the changes observed with the latter two classes of amino acids, the pD dependence of the cytosine spectrum was carefully studied. The first set of data on this is presented in the upper part of Table II and is plotted in Figs. 3 and 4. Thus, certainly significant changes of  $H_5$  and  $H_6$  chemical shifts (and the  $J_{56}$  coupling) are affected by pD changes. In addition, some rather significant changes of line widths with pD occur for both  $H_5$  and  $H_6$ . To show that these line width changes

TABLE I - Chemical Shifts For  $H_5$  and  $H_6$  and  $J_{56}$

For Cytosine Admixed With Various Amino Acids.

The pD Of These Solutions Was Not Determined.

	Chemical Shifts			Coupling Constants		Molar Amino Acid Conc.
	$H_6$	$H_5$	$H_6-H_5$	${}^6J_{56}$	${}^5J_{56}$	
Cytosine	429.8	276.4	153.4	7.2	7.2	
Glycine	430.8	277.4	153.4	7.1	7.2	.84
L-Alanine	431.1	277.5	153.6	7.0	7.2	1.1
L-Lysine	432.4	278.4	154.0	7.2	7.1	1.0
L-Leucine	430.2	277.2	153.0	7.1	7.0	.17
DL-Isoleucine	430.2	276.7	153.5	7.3	7.2	.17
DL-Valine	431.0	277.6	153.4	7.2	7.2	.69
DL-Proline	431.4	277.4	154.0	7.3	7.2	1.1
L-Phenylalanine	428.2	275.6	152.6	7.2	7.2	.12
L-Serine	431.5	277.9	153.6	7.2	7.2	.43
L-Glutamic Acid	443.8	287.5	156.3	7.3	7.4	.045
L-Cysteine	433.4	279.4	154.0	7.2	7.2	.94
L-Histidine	431.2	278.0	153.2	7.2	7.2	.26
L-Threonine	430.2	277.0	153.2	7.2	7.2	.13
L-Methionine	430.5	277.2	153.3	7.2	7.2	.20
L-Glutamine	431.1	277.5	153.6	7.2	7.2	.25
L-Aspartic Acid	441.3	285.6	155.8	7.4	7.4	.033
L-Arginine·HCl	431.7	278.2	153.6	7.1	7.0	.50
L-Asparagine·H <sub>2</sub> O	430.9	277.4	153.5	7.2	7.3	.17
L-Tryptophan	426.6	274.0	152.5	7.2	7.2	.055
L-Hydroxyproline	431.9	278.1	153.8	7.1	7.2	.92
L-Tyrosine Basic*	461.01	273.36	187.65	6.9	6.9	.16
L-Cystine Basic*	456.47	268.90	187.57	6.9	6.9	.15
Basic Cytosine*	457.19	269.08	188.11	6.9	6.9	

\* Solution of these insoluble amino acids was achieved by adding 2 drops 40% NaOD; pD probably about 14; see Table II; the same amount of NaOD was added to the stock cytosine solution.

Fig. 1. Cytosine H<sub>6</sub> proton chemical shift with added amino acid; pD's unknown.



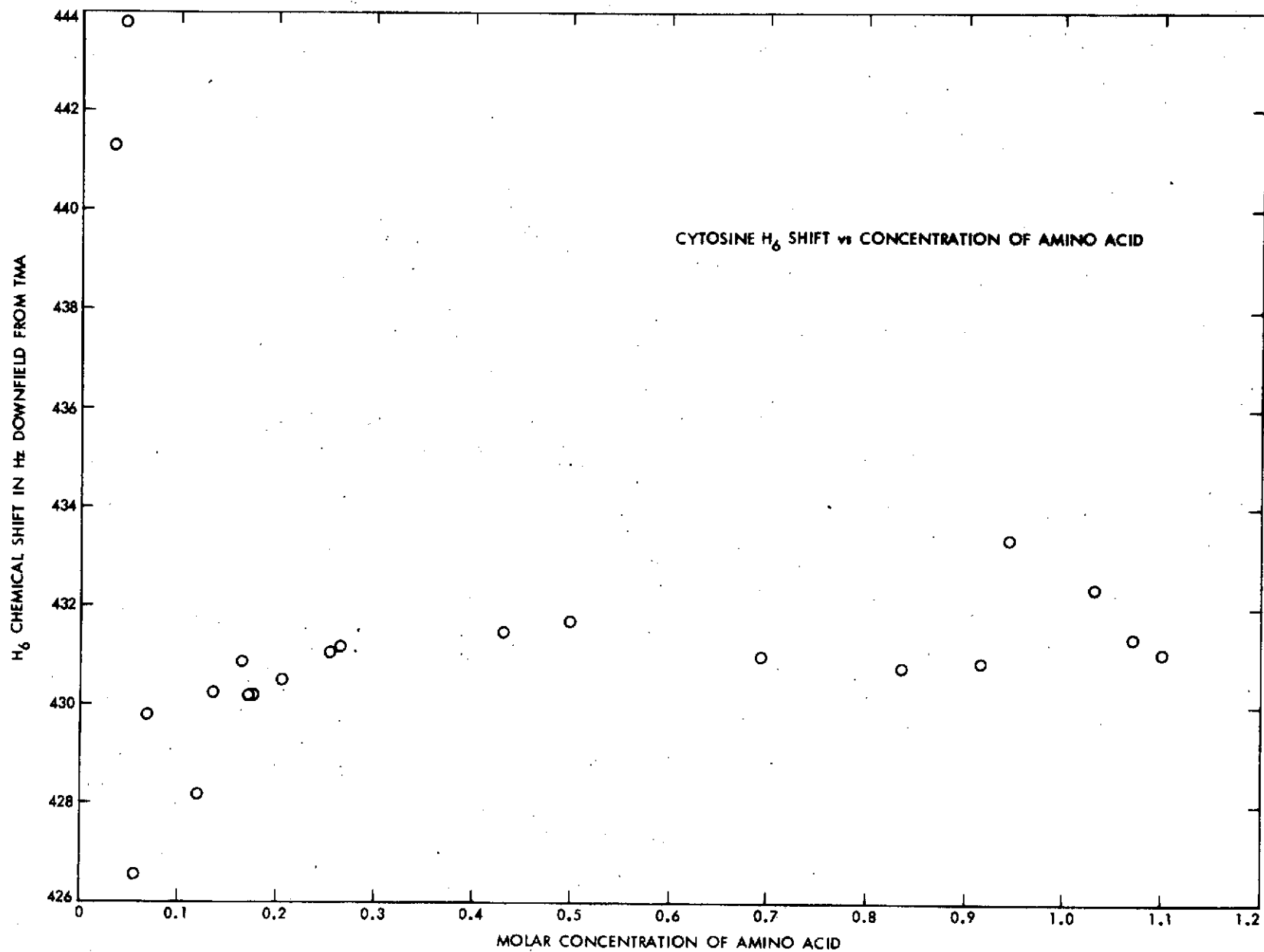


Fig. 2. Cytosine H<sub>5</sub> proton chemical shift with added amino acid;  
pD's unknown.

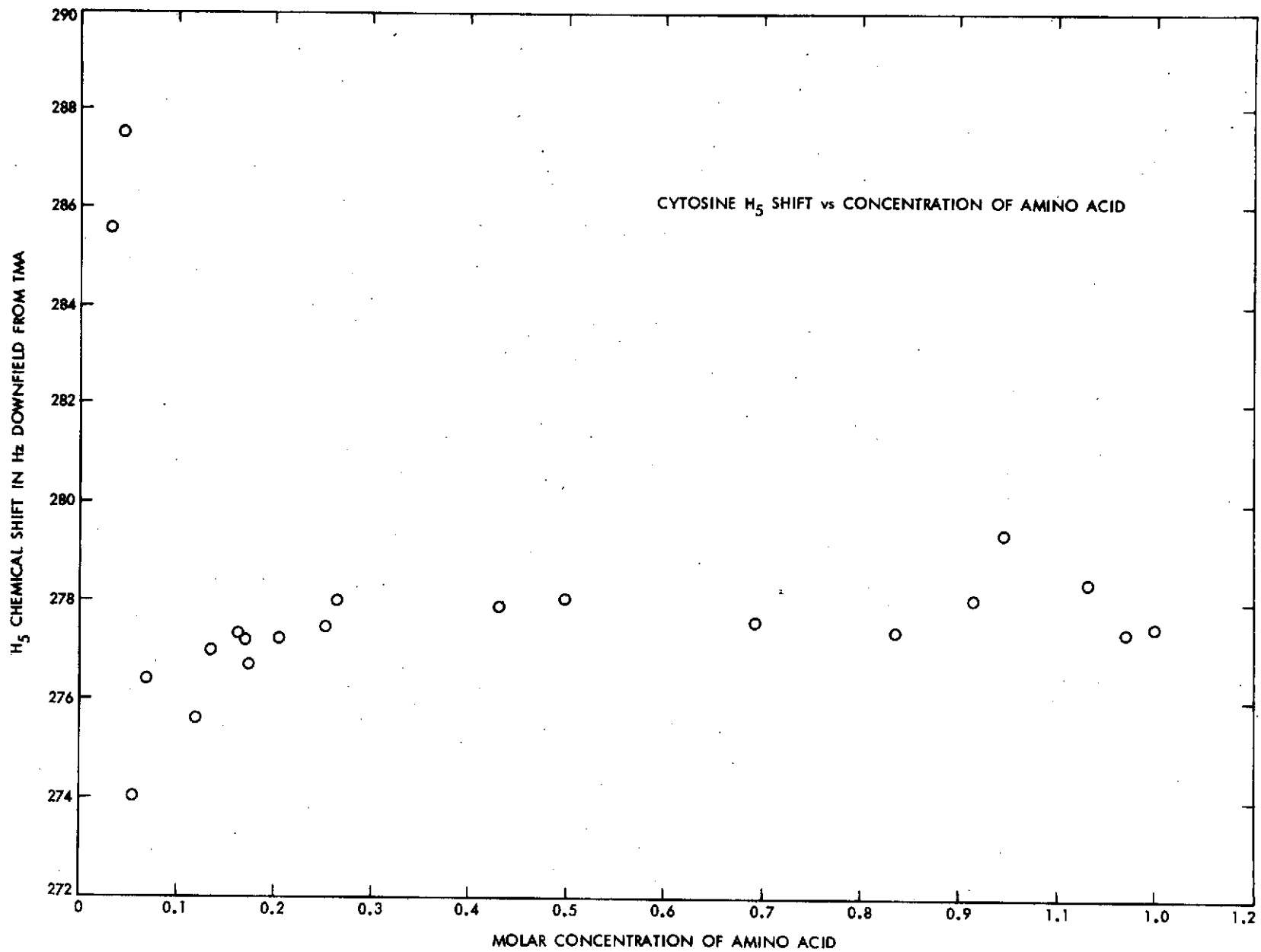


TABLE II - Chemical Shift, Coupling Constant And Line Width  
For Cytosine In D<sub>2</sub>O At Various pD's;

Chemical Shifts Referenced Internally To TMA; All Parameters In Hzs.

A Cytosine Before Celex-100 Treatment

pD-0.40	1.08	2.80	5.21	7.33	8.81	8.95	11.33	13.58
H <sub>6</sub>	454.72	454.10	438.56	430.94	430.85	430.59	431.09	455.86
H <sub>5</sub>	296.08	295.36	283.87	278.08	277.84	277.83	277.73	268.27
<sup>6</sup> J <sub>56</sub>	7.56	7.58	7.24	7.20	7.14	7.20	7.10	5.86
<sup>5</sup> J <sub>56</sub>	7.54	7.54	7.2	7.14	7.16	7.08	7.04	5.78
(H <sub>6</sub> -H <sub>5</sub> )	158.49	158.74	154.69	152.86	153.01	152.76	153.36	187.59
Δω <sub>5</sub>	.5	.5	.6	.6	.6	.6	.9	1.3
Δω <sub>6</sub>	.6	.6	2.5	1.7	.8	.8	.7	.6

B. Cytosine After Celex-100 Treatment

pD-0.40	1.19	3.33	5.07	6.81	8.54	10.08	11.81	13.66
H <sub>6</sub>	454.57	453.63	440.89	430.81	430.69	430.61	432.53	458.70
H <sub>5</sub>	295.75	294.89	285.34	277.91	277.73	277.79	277.30	272.48
<sup>6</sup> J <sub>56</sub>	7.56	7.54	7.30	7.18	7.22	7.14	7.14	5.92
<sup>5</sup> J <sub>56</sub>	7.62	7.54	7.4	7.26	7.22	7.18	7.12	5.88
(H <sub>6</sub> -H <sub>5</sub> )	158.82	158.74	155.55	152.90	152.96	152.82	155.23	186.22
Δω <sub>5</sub>	.5	.5	.5(.5)*	.5	.6	.7	1.6(0.6)*	1.0(.6)*
Δω <sub>6</sub>	.5	.7	3.5(.8)*	.8	.9	1.0	1.0(0.6)*	.6(.6)*

\*Line width after repeated treatment with Dithizone-CCl<sub>4</sub> solution then spectra recorded with small drop of later solution in bottom of NMR tube.

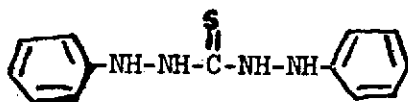
result from the presence of paramagnetic ion impurities in the solutes, a stock solution of cytosine was prepared which was treated with Celex-100 to remove these ions. The set of results given in the lower part of Table II suggest that the removal of paramagnetic ion was not successful. However, it was discovered that this apparent result was due to paramagnetic ion impurities associated with the new NMR tubes used in this work. Normal procedure was to wash all new NMR tubes with distilled water known to have insignificant paramagnetic ions and then subsequently wash the tubes with purified acetone and dry. Although the NMR spectra of cytosine solutions in these tubes both before and after Celex-100 treatment showed broadened spectra at certain pD's, on addition of amino acid, much of the broadening disappeared. It was later found that if the tubes were washed with strong acid, then distilled water and then acetone followed with drying, no broadening occurred.

A rather convenient procedure was found by which the line broadening existing in the previously prepared samples was essentially eliminated. A solution of about  $10^{-6}$  M dithizone, whose structure is depicted below, was prepared.<sup>44</sup> A few drops of this solution was added to an already

---

44. For some comment and refs. regarding the use and properties of dithizone for complexing metal ions see: A. J. Bauman and J. H. Richards, Sep. Sci., 6, 715 (1971).

---



Dithizone

prepared  $D_2O$  solution of cytosine in a 5 mm O.D. NMR tube. The tube was agitated, and it was noted that the color of the dithizone solution faded. The latter was removed with a fine dropper and several more drops were added, followed by further agitation. This treatment was repeated until the dithizone solution no longer showed evidence of color change upon agitation. Spectra of solutions thus treated showed significantly less line broadening. However, a return of line broadening was noted on standing for these samples, thus suggesting that paramagnetic species were being slowly released from the surface of the NMR tubes. It was found that this return of line broadening could be stopped by leaving a drop of the  $CCl_4$ -dithizone solution in the bottom of the NMR tube containing the  $D_2O$  solution of cytosine or other nucleic acid constituent. No loss of spectral resolution was noted for such two phase solutions for cytosine, adenine and adenosine-5'-monophosphate (5'-AMP) solutions. The line widths for cytosine solutions treated this way are shown in parentheses in the lower part of Table II. No change in chemical shift was noted for the  $H_5$  and  $H_6$  protons on removal of the paramagnetic ion by the treatment described above.

The pD dependence of the  $H_5$  and  $H_6$  proton chemical shifts are shown in Figs. 3 and 4. The observed curves show changes and inflexion points corresponding approximately to the pK's of cytosine ( $pK_1 = 4.60$  and  $pK_2 = 12.16$ ).<sup>45</sup> An interesting, heretofore unreported, observation was that at

---

45. D. O. Jordan in "The Nucleic Acids," edited by E. Chargoff and J. N. Davidson, Academic Press, Inc., N.Y., Vol. I, pp. 455-458.

---

Fig. 3. Cytosine H<sub>6</sub> proton chemical shift pD profile.

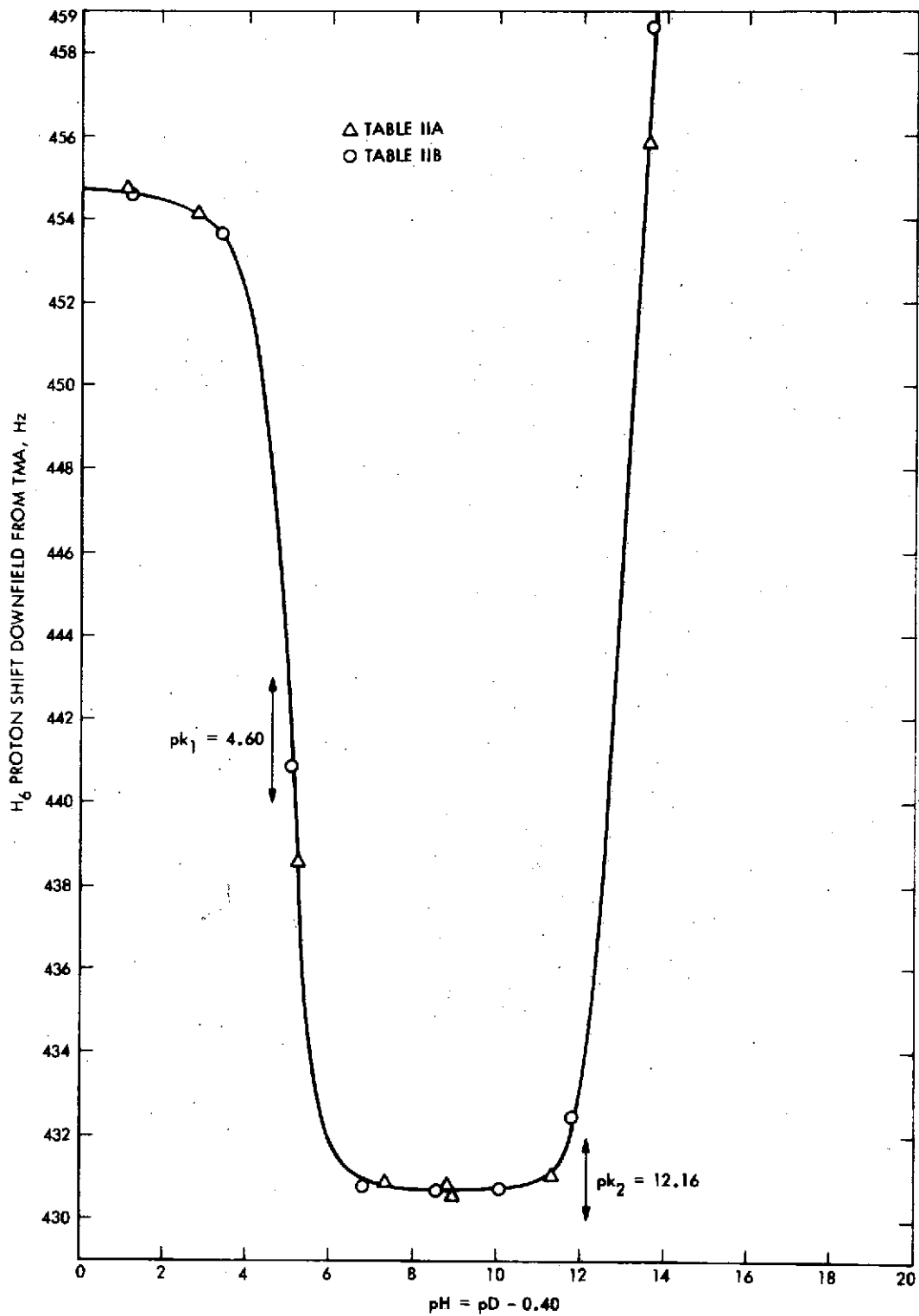
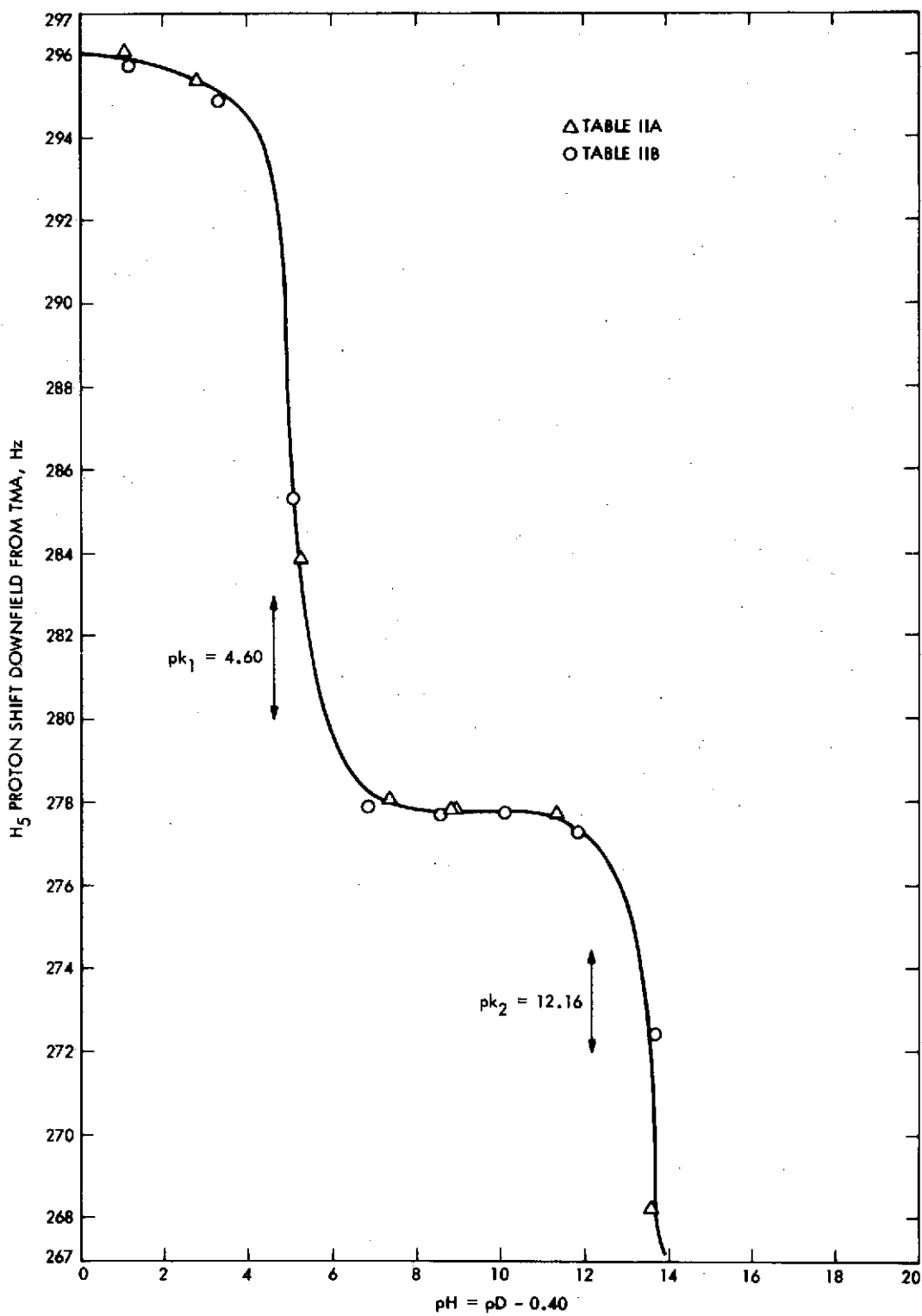




Fig. 4. Cytosine H<sub>5</sub> proton chemical shift pD profile.



very high pD the H<sub>6</sub> proton was deshielded while the H<sub>5</sub> proton is increasingly shielded. In addition, the chemical shift of H<sub>6</sub> seems to be most sensitive to pH changes. Thus, the change in pH from 1 to about 9 causes a 24 Hz change for H<sub>6</sub> while only a 18 Hz for H<sub>5</sub>. The change from pD 9 to 13.6 causes a 27.0 Hz downfield shift for H<sub>6</sub> and a 8 Hz upfield shift for H<sub>5</sub>. The coupling J<sub>56</sub> also is sensitive to pD. It is about 7.6 Hz at low pD, 7.2 Hz around neutral and drops to about 5.8 Hz at pD 13.6. These data are plotted in Fig. 5. Complete understanding of these observed changes, however, were not sought as they were outside the scope of this work. It was felt that initially the studies to be accomplished here would be concerned with pH's and pD's close to neutral. Whether this is realistic in regard to primitive earth conditions is certainly open to question.

In light of the changes apparent in Figs. 1-5, cytosine solutions adjusted close to pD 7.4 were studied for one acidic amino acid, glutamic acid, and one aromatic acid, tryptophan. Both of these amino acids are rather insoluble but it is apparent from the data shown in Table III that they cause measureable changes in the H<sub>5</sub> and H<sub>6</sub> chemical shifts and possibly in the case of tryptophan a just detectable diminution in J<sub>56</sub>. Added glutamic acid appears not to affect the H<sub>5</sub> chemical shift measureably but causes a decrease of the H<sub>6</sub> chemical shift. A possible cause for this change probably arises from some interaction closer to H<sub>6</sub> than H<sub>5</sub>. A possible mechanism would be hydrogen bonding interaction involving the amino nitrogen proton which could result in inductive reduction of the electron density at C-6. The shielding effect observed with tryptophan

Fig. 5. Cytosine J<sub>56</sub> pD profile.

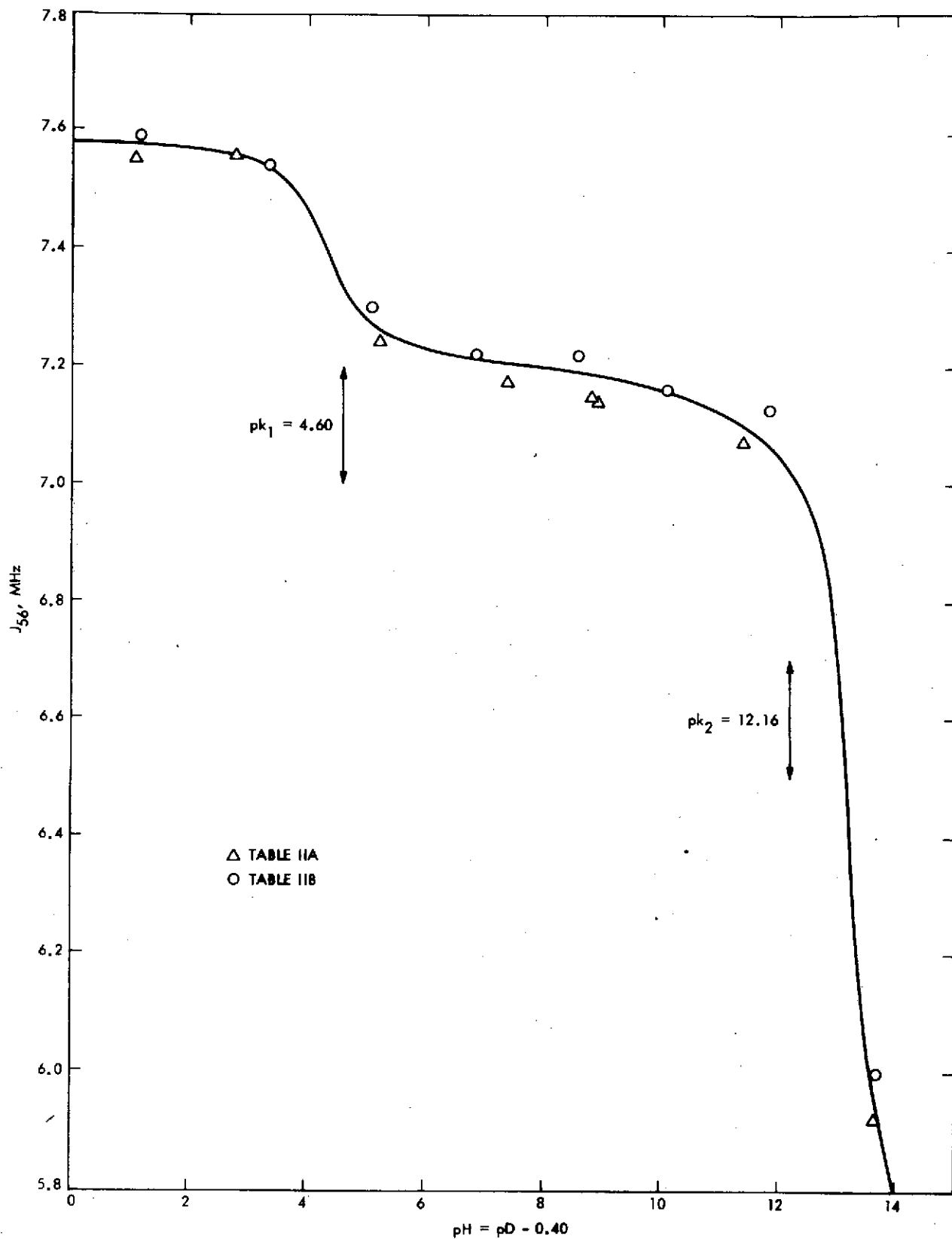
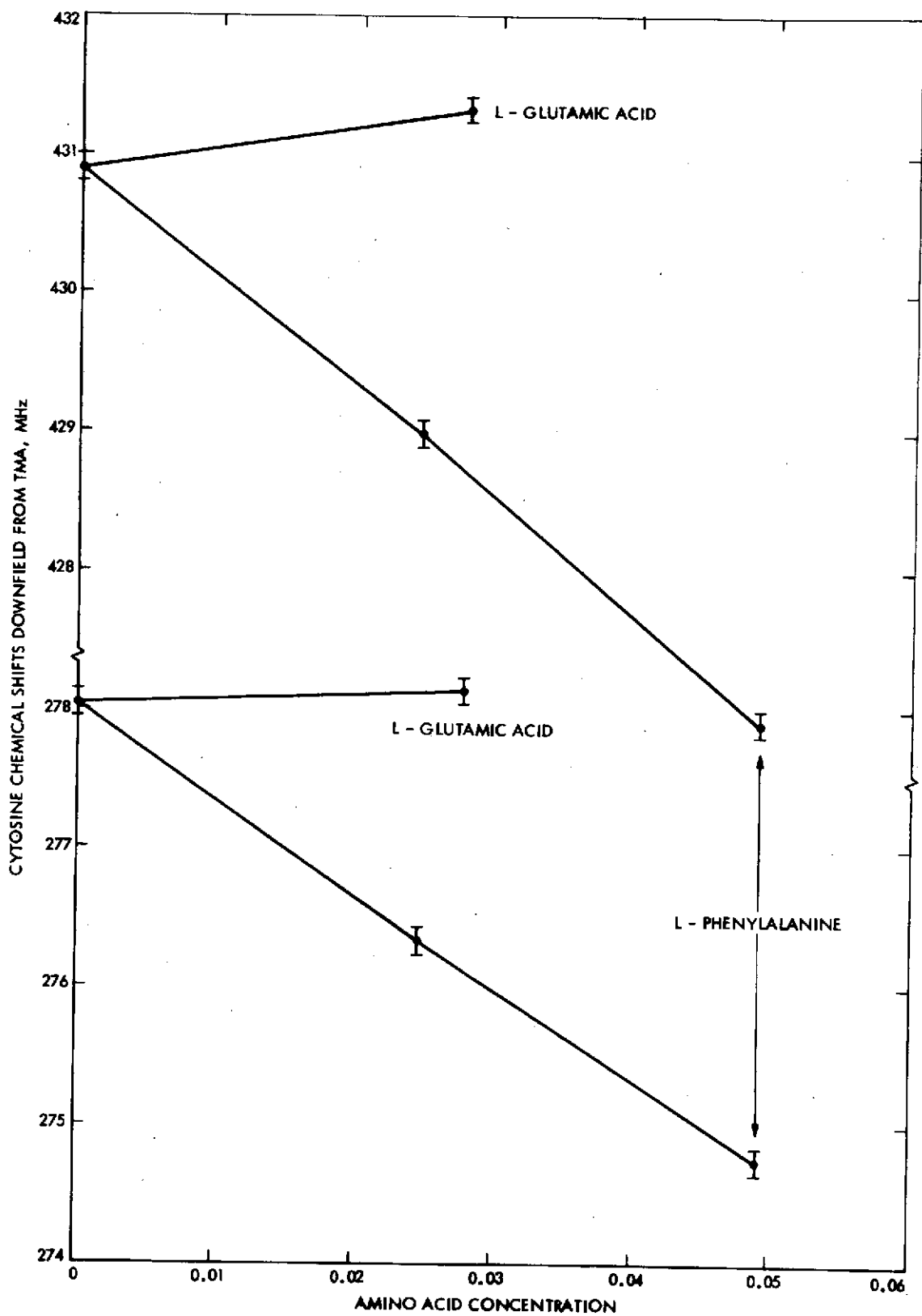


TABLE III - Cytosine Chemical Shifts and Coupling Constants  
In Hz At pD 7 With Added Amino Acid

	Molar Amino Acid Conc.	Chemical Shifts				
		pD-0.40	H <sub>6</sub>	H <sub>5</sub>	(H <sub>6</sub> -H <sub>5</sub> )	J <sub>56</sub>
Cytosine Only (.0279 M)	-----	7.33	430.98	278.15	152.83	7.17
	-----	7.00	430.90*	278.04*	152.86*	----
	-----	6.81	430.81	277.91	152.90	7.22
Plus <u>L</u> -Glutamic Acid	.0276	7.12	431.33	278.13	153.20	7.20
Plus <u>L</u> -Tryptophan	.0247	7.00	428.98	276.33	152.65	7.14
	.0493	7.11	426.90	274.74	152.16	7.12

\* Extrapolated from pH (observed) 7.33 and 6.81 values.

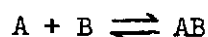
Fig. 6. Cytosine H<sub>5</sub> and H<sub>6</sub> proton chemical shifts in presence of glutamic acid and phenylalanine.





most certainly results from intercalation between aromatic rings. Fig. 5 plots the changes with concentration. The changes, if due to stacking, seem to be comparable to those observed for tryptophan and tryptamine with other nucleic acid derivatives.<sup>31-33</sup> Careful measurement of the tryptophan and glutamic acid protons chemical shifts was not made.

If one assumes that the changes of the H<sub>6</sub> and H<sub>5</sub> chemical shifts with changes of amino acid concentration at constant cytosine concentration arise from 1:1 complex formation, it is possible to estimate the association constants approximately even for the limited data presented here. Consider the following association equation:



Where the rate is fast on the NMR time scale, it can be shown that

$$\frac{\delta - \delta_A}{\delta_{AB} - \delta_A} = \frac{[AB]}{[A_0]} = \frac{\Delta}{\Delta_0} \quad (1)$$

where  $\delta$  is the shift observed,  $\delta_A$  is the chemical shift of free A,

$\delta_{AB}$  is the shift of A in the complex,  $[AB]$  the concentration of the complex,  $[A_0]$  the initial A concentration,  $\Delta = \delta - \delta_A$  and  $\Delta_0 = \delta_{AB} - \delta_A$

Now,

$$K = \frac{[AB]}{[A][B]} \quad (2)$$

for a 1:1 complex where K is the equilibrium constant and  $[A]$  and  $[B]$  are the equilibrium concentrations. Combining Eqs. 1 and 2 gives

$$\frac{\Delta}{\Delta_0} = \frac{K[A][B]}{[A_0]} \quad (3)$$

There are a number of ways Eq. 3 may be treated, but most assume either  $[A] \gg [B]$  or  $[B] \gg [A]$ .<sup>32,46</sup> This is not our case as

---

46. See for example, R. Foster and C. A. Fyfe in "Progress in N.M.R. Spectroscopy", edited by J. W. Emsley, J. Feeney and L. H. Sutcliffe, Pergamon Press, N.Y., Vol. 4, pp. 1-89.

---

$[A_0] = 0.0279$  M and  $[B_0]$ 's are 0.0493 and 0.0247 M. In this case the most valid procedure would be to iterate for various values of K and  $\Delta_0$  but this would only be justified if more data were available. There is some data for complexing of cytosine with purine in which  $\Delta_0$ 's were determined to be 54 Hz and  $H_5$  and 39 Hz for  $H_6$ .<sup>32</sup> Equation 3 can be changed to

$$\frac{\Delta [A_0]}{\Delta_0 [A][B]} = K. \quad (4)$$

Thus, if we knew the  $\Delta$ 's for both protons at one set of concentrations, then

$$\frac{\Delta^5}{\Delta_0^5} = \frac{\Delta^6}{\Delta_0^6} \quad \text{or} \quad \frac{\Delta_0^6}{\Delta_0^5} = \frac{\Delta^6}{\Delta^5} \quad (5)$$

where  $\Delta_0^5$  and  $\Delta_0^6$  correspond to the  $\Delta_0$  definition given above and  $\Delta^5$  and  $\Delta^6$  are measured at a particular concentration. The ratio for our case appears to be about 1:2, that is  $\Delta_0^6 > \Delta_0^5$  in contrast to the case of complexation with purine. Since there is evidence that tryptophan may cause chemical shift changes similar to those of purine on complexing, to get a rough estimate of K for cytosine, we have calculated K's for  $\Delta_0^6$  at 40, 50 and 60 Hz which yield K's of 2.0, 1.6 and 1.3 M<sup>-1</sup> which are reasonable.<sup>32</sup> These were gotten from Eq. 4 assuming  $[A_0] \approx [A]$  and  $[B_0] \approx [B]$ .

This approximate estimate is not far off if one uses the estimated K's to calculate  $[A]$  or  $[B]$ .

The chemical shifts of the  $H_2$  and  $H_8$  adenine protons were next studied in the presence of added amino acid. This study was limited to selected members of the several classes of amino acids: L-valine for an aliphatic one, L-glutamic acid for an acidic one, L-phenylalanine as an aromatic one, L-lysine as a basic one and L-arginine as a basic one, but more important as one for which evidence exists that its polymers interact with nucleic acid constituents. Few studies have been made with adenine because of its low solubility in water. It is our opinion that because of the small chemical shift between the  $H_2$  and  $H_8$  protons that a definitive spectral assignment of which proton is which has not been made. However, for our purposes an exact assignment is not necessary as long as we are alert to the possibility that a chemical shift inversion does not take place. Table IV and Figs. 7 and 8 summarize the data on the  $H_2$  and  $H_8$  chemical shifts from these experiments with adenine. Several types of behavior are readily apparent. We have tentatively taken  $H_2$  as downfield from  $H_8$  in analogy with purine. That this is probably the correct assignment will be discussed below. In the presence of phenylalanine an apparent linear upfield shift with concentration occurs with both protons. This is most probably due to base stacking involving adenine and the amino acid phenyl ring. In contrast, an apparent linear downfield shift with glutamic acid concentration is evident. This deshielding cannot all be due to destacking of the adenine or ionic strength effects as significantly larger concentrations of the three more soluble amino acids cause a lesser effect per unit concentration. Although more evidence is needed, it is not

TABLE IV - Chemical Shifts Of Adenine H<sub>2</sub> and H<sub>8</sub> Protons In Half Saturated Solution In D<sub>2</sub>O As A Function Of Added Amino Acids; Internal Standard Is TMA

	Chemical Shifts					
	Molar Concn. pD-0.40		H <sub>2</sub> <sup>a</sup>	H <sub>8</sub> <sup>a</sup>	H <sub>2</sub> -H <sub>8</sub> <sup>b</sup>	$\Delta_{28}$ <sup>c</sup>
Half Saturated Solution <sup>c</sup>	0.0040	7.29	502.08	497.10	5.02	----
<u>L</u> -Valine	0.502	7.12	504.10	499.86	4.20	0.82
	0.0999	7.05	503.60	498.74	4.86	0.16
<u>L</u> -Lysine	0.500	7.09	503.24	500.38	2.86	2.16
	0.101	7.14	503.16	498.44	4.72	0.30
<u>L</u> -Arginine	0.501	6.99	503.10	500.98	2.24	2.78
	0.101	6.99	503.60	499.34	3.90	1.12
<u>L</u> -Glutamic Acid	0.0342	6.99	504.46	499.20	5.26	-0.24
	0.0167	7.02	503.20	498.10	5.10	-0.08
<u>L</u> -Phenylalanine	0.165	7.12	498.62	494.56	3.95	1.07
	0.0822	7.10	500.44	496.24	4.30	0.72
Saturated Solution <sup>c</sup>	0.0080	7.10	500.87	496.44	4.43	----
Extrapolated To Zero Concentration	-----	----	507.7	502.1	5.6	----

a) Hz downfield from center of TMA resonance.

b) In most cases this spacing was obtained from the chemical shifts given in the two columns to the left but it was found that where the resolution, saturation and sweep rates were not optimum, a more accurate value of this difference could be extracted from observed spectra by measuring the distance between the first observed ring after each peak. This was especially true when the fast passage line widths were of the order of 2 Hz wide.

c) Concentrations determined by NMR

Fig. 7. Adenine proton chemical shifts in presence of amino acids.

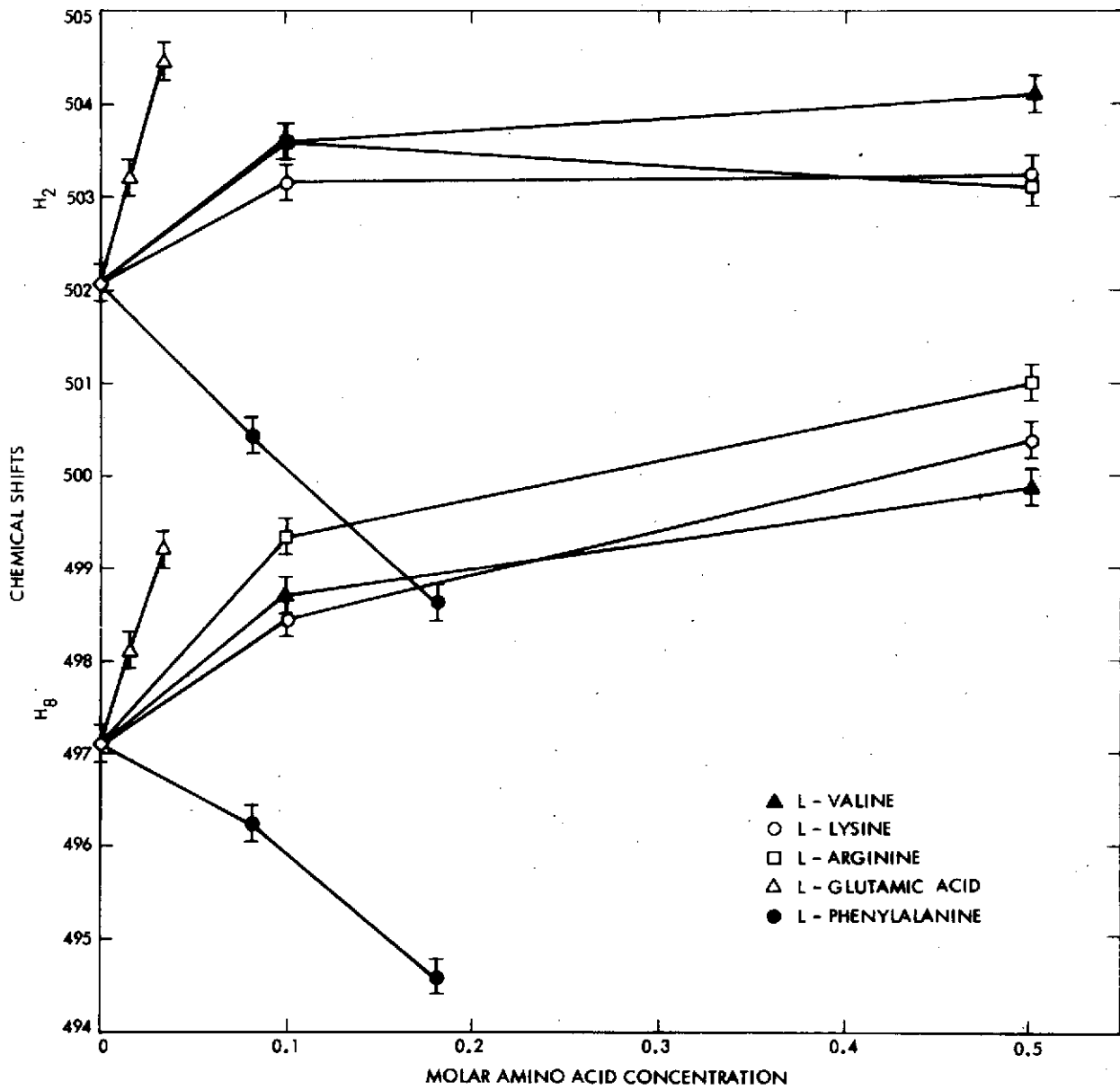
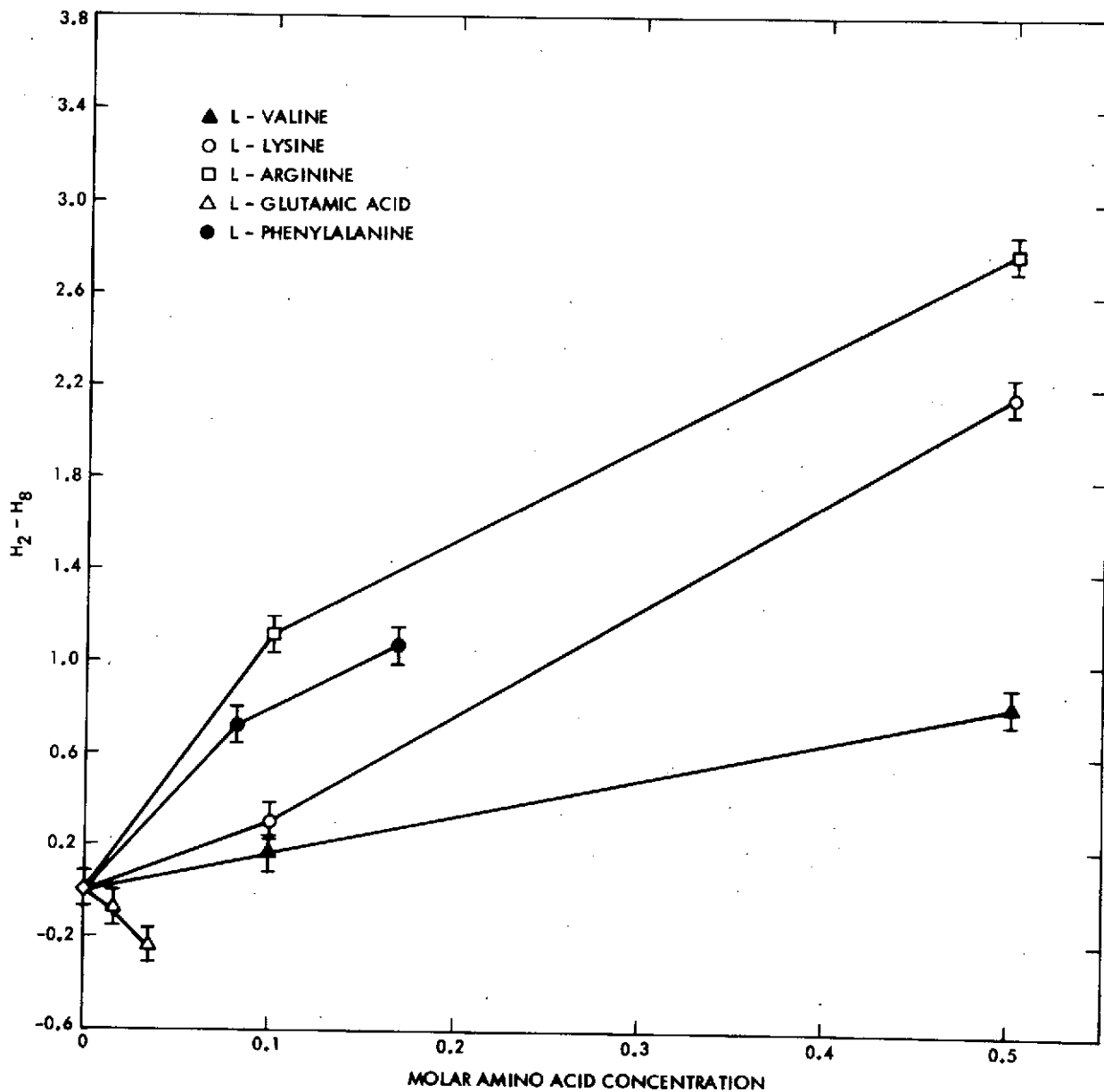


Fig. 8. Adenine H<sub>2</sub> - H<sub>8</sub> spacing in presence of amino acids.





impossible that glutamic acid makes some specific hydrogen bonded complex with adenine. The three soluble amino acid appear to deshield the adenine H<sub>2</sub> and H<sub>8</sub> protons significantly less per unit concentration than glutamic acid or the shielding caused by phenylalanine. It is interesting to consider the difference between the H<sub>2</sub> and H<sub>8</sub> chemical shifts for any particular solution. Figure 8 exhibits the data plotted this way. Four of the amino acids cause this spacing to decrease significantly while in the case of glutamic acid this spacing appears to increase.

In Table IV are data for adenine chemical shifts at two concentrations. The observed differences stem from its self-association by vertical stacking, as has been studied in detail for purine and a number of other purine derivatives by NMR<sup>32,47</sup> and other techniques.<sup>48,49</sup> To determine

---

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

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

49. P. O. P. Ts'o, I. S. Melvin and A. C. Olson, J. Am. Chem. Soc., 85, 1289 (1963).

---

the self-association constant, it is necessary to determine the  $\delta_A$  which is the chemical shift at zero concentration for both protons. With  $\delta_A$  known, the following equation can be used to get the self-association constant for dimer formation:<sup>32,46</sup>

$$\left(\frac{\Delta}{[A_0]}\right)^{\frac{1}{2}} = \left(\frac{2K}{\Delta_0}\right)^{\frac{1}{2}} (\Delta_0 - \Delta) \quad (6)$$

In the case of n-mer formation, the equation which applies<sup>32</sup> is

$$\left(\frac{\Delta}{[A_0]}\right)^{\frac{1}{2}} = \left(\frac{K}{2\Delta_0}\right)^{\frac{1}{2}} (2\Delta_0 - \Delta) \quad (7)$$

For dimer formation, a plot of  $(\Delta/[A_0])^{\frac{1}{2}}$  against  $\Delta$  will give a straight line with slope  $(2K/\Delta_0)^{\frac{1}{2}}$  and X-axis intercept  $\Delta_0$ . For n-mer formation, the slope is then  $(K/2\Delta_0)^{\frac{1}{2}}$  and the intercept is  $2\Delta_0$ . With the two data points of the present work, it is not possible to determine  $\delta_A$  using Eq. 6 in the normal way. However, it is possible to get a rough estimate of K for adenine in the following way. If Eq. 6 is squared and manipulated slightly, one can solve for  $\Delta$ :

$$\Delta = \frac{1 + 4K[A_0] \pm \sqrt{8K[A_0] + 1}}{2C} \quad (8)$$

where  $C = 2K[A_0]/\Delta_0$ . Expanding and using the minus sign above yields

$$\Delta = 2K[A_0]\Delta_0 (1 - 4K[A_0] + \dots) \quad (9)$$

Recalling that  $\Delta = \delta - \delta_A$  gives

$$\delta_A = \delta - 2K[A_0]\Delta_0 (1 - 4K[A_0] + \dots) \quad (10)$$

Now if we have values of  $\delta$  at several  $[A_0]$ 's for two protons, we can estimate the ratio of  $\Delta_0$ 's for these protons because

$$\frac{{}^1\delta_0 - {}^2\delta_0}{{}^1\delta_v - {}^2\delta_v} = \frac{{}^D\Delta_0}{{}^U\Delta_0} \quad (11)$$

where  $^1\delta_D$  and  $^1\delta_U$  are the chemical shifts for, in our case of adenine, the downfield proton and upfield proton at a particular  $[A_0]$ . Correspondingly,  $^2\delta_D$  and  $^2\delta_U$  are the chemical shifts at another  $[A_0]$ . This ratio appears to be about 1:2; thus  $^D\Delta_0 > ^U\Delta_0$ . In purine<sup>32,47</sup> and 6-methylpurine<sup>47</sup> the  $H_2$  protons appear to have the largest  $\Delta_0$ , so on this basis the downfield proton in adenine can be assigned to  $H_2$ . An estimate of  $^D\Delta_0$  as about 50 Hz is not unreasonable based on previous results for purine<sup>32,47</sup> and 6-methylpurine.<sup>47</sup> With this value and from the fact that

$$\Delta[A_0] = 2K[A]^2\Delta_0, \quad (12)$$

an approximate self-association constant can be found if we have  $\Delta$ .

Thus,

$$K \approx \frac{\Delta}{2[A_0]\Delta_0} \quad (13)$$

where also  $[A]$  can be reasonably assumed to be nearly  $[A_0]$  for the low concentration at which the data were taken. Because  $\Delta = \delta - \delta_A$  and we only have  $\delta$ , we can at best use a linear extrapolation to obtain a  $\delta_A$  which then gives a lower limit for  $K$  of about  $8 M^{-1}$ . This  $K$  is the same order as has been observed for another 6-substituted purine, 6-methylpurine.<sup>48</sup>

The ratio  $^D\Delta_0 / ^U\Delta_0$  mentioned above was taken to indicate that  $H_D = H_2$  and  $H_U = H_8$ . Further confirmation of this assignment has come from the fact that several adenine  $D_2O$  samples, prepared early in this work and heated for various periods to affect solution, exhibited substantial reduction in the intensity of the upfield proton signal. It is well known that  $H_8$  of purines exchanges with  $D_2O$ .<sup>47</sup>

The next sets of experiments which were carried out were spectral studies of adenonine-5'-monophosphate (5'-AMP) mixed with the same five amino acids with which adenine was studied. It should be realized that studies with the monophosphate might involve introduction of too many complications and perhaps similar studies with the nucleoside should have been carried out first. Such studies were planned, but due to limitation of the scope of this work they could never be carried out.

Spectra were recorded for two 5'-AMP concentrations at two different amino acid concentrations. The data are summarized in Table V and plotted in Figs. 9, 10 and 11. From these data and some for 5'-AMP at 0.49 M concentration, the self-association constants to dimers for the latter were estimated by using Eq. 6. Thus,  $K_2 \approx 2$ ,  $K_8 \approx 3$  and  $K_{1'} \approx 1 \text{ M}^{-1}$  and  ${}^2\Delta_o \approx 40 \text{ Hz}$ ,  ${}^8\Delta_o \approx 53 \text{ Hz}$  and  ${}^{1'}\Delta_o \approx 48 \text{ Hz}$  were found for dimer formation. No similar set of parameters appear to have been published heretofore. The K's are definitely less than those estimated for adenine. The differences in K's reflect the fact that 5'-AMP possesses several sites involved in the self-association process.<sup>32</sup>

Because of the multifunctionality of 5'-AMP, several types of interaction sites are expected and the limited data (only two amino acid and two 5'-AMP concentrations) taken in the present work, the data will only be discussed qualitatively. Consideration of multisite complexing can be quite complicated.<sup>32</sup>

For a 5'-AMP concentration of 0.01 M, phenylalanine appears to cause an increase of shielding similar in degree to that observed for adenine. However, at a 5'-AMP concentration of .05 M, although the  $H_2$  and  $H_{1'}$  protons exhibit an increase of shielding, the  $H_8$  shows a decrease over that at 0.01 M

TABLE V - Chemical Shifts of  $H_2$  and  $H_8$  of Adenosine-5-Monophosphate in  $D_2O$  As A Function Of Added Amino Acid

	Molar Concen.	Obs. pH	$H_8$	$H_2$	$H_{1'}$	$H_8-H_2$	$J_{1'2}$
5'-AMP 0.01M	-----	6.98	536.5	506.5	295.2	30.0	6.0
<u>L</u> -Valine	0.500	6.93	536.5	506.7	295.0	29.8	5.8
	0.102	6.03	535.9	506.5	295.1	29.4	5.8
<u>L</u> -Lysine HCl	0.500	6.93	537.5	505.8	293.4	31.7	5.5
	0.100	7.10	538.1	505.7	294.2	33.4	5.9
<u>L</u> -Arginine HCl	0.500	6.97	538.0	505.9	292.4	32.1	5.8
	0.100	7.04	538.1	505.7	294.3	32.4	5.8
<u>L</u> -Glutamic Acid	0.032	7.00	536.6	506.1	295.2	30.5	5.8
	0.016	7.02	536.5	506.4	295.3	30.1	6.0
<u>L</u> -Phenylalanine	0.163	6.95	535.0	502.6	292.2	32.4	5.8
	0.082	7.11	535.8	503.1	292.4	32.7	5.9
5'-AMP 0.05M		6.95	533.3	499.2	292.3	34.1	5.8
<u>L</u> -Valine	0.503	6.95	535.2	501.4	292.8	33.8	5.7
	0.101	6.99	534.4	500.6	292.8	33.8	5.9
<u>L</u> -Lysine HCl	0.500	6.95	534.5	499.1	290.5	35.4	5.5
	0.100	6.79	532.7	500.0	291.9	32.7	5.6
<u>L</u> -Arginine HCl	0.500	6.84	534.8	500.9	290.2	33.9	5.6
	0.101	6.82	533.1	499.5	291.3	33.6	5.8
<u>L</u> -Glutamic Acid	0.326	6.95	534.3	501.6	293.1	32.7	5.8
	0.017	6.79	531.7	499.9	292.4	31.8	5.5
<u>L</u> -Phenylalanine	0.164 <sup>a</sup>	7.20	534.3 <sup>c</sup>	496.6 <sup>c</sup>	289.0 <sup>c</sup>	37.7	5.6
	0.082 <sup>b</sup>	6.98	532.8 <sup>c</sup>	498.2 <sup>c</sup>	290.7 <sup>c</sup>	34.6	5.6
5'-AMP 0.49 M	-----	6.84	518.0	479.0	278.5	39.0	5.0

a) 5'-AMP - concentration 0.0429

b) 5'-AMP - Concentration 0.0427

c) Chemical in absence of amino acid at those concentrations  $H_2$  533.8,  
 $H_8$  500.3,  $H_{1'}$  292.8 and  $H_8-H_2$  33.5.

Fig. 9. Chemical shift of  $H_8$  of 5'-AMP in presence of amino acids.

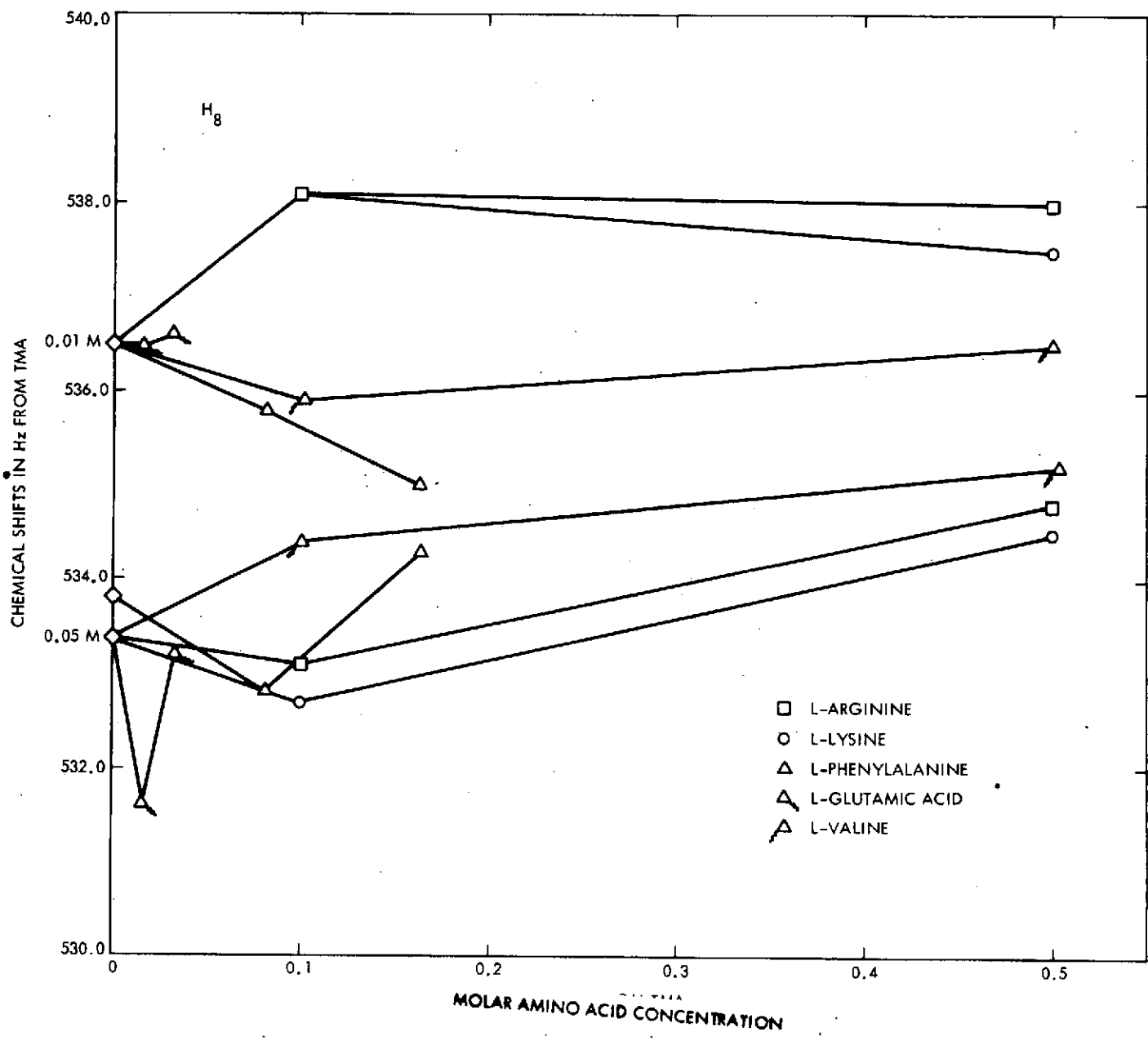


Fig. 10. Chemical shifts of  $H_2$  of 5'-AMP in presence of amino acids.



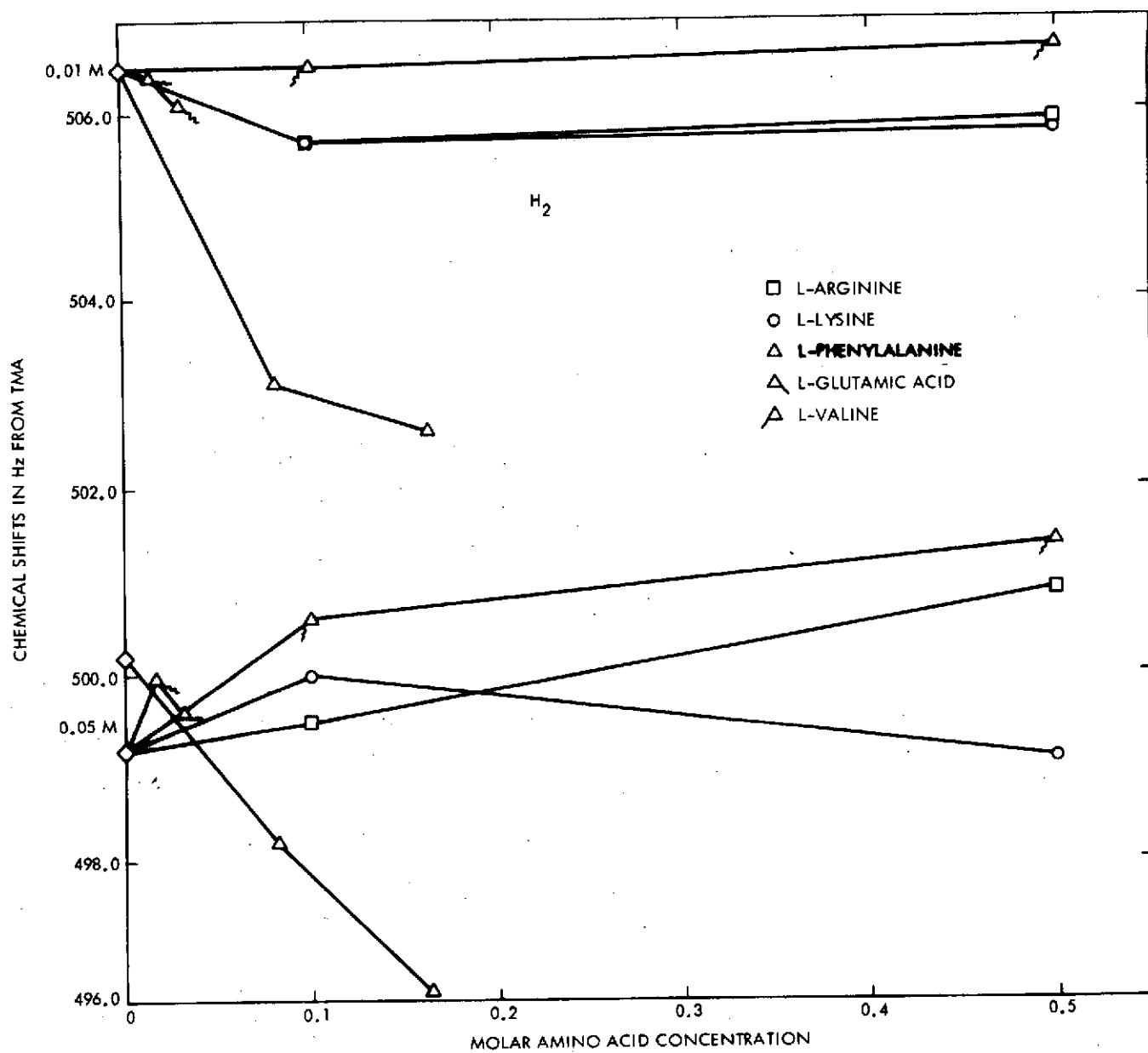
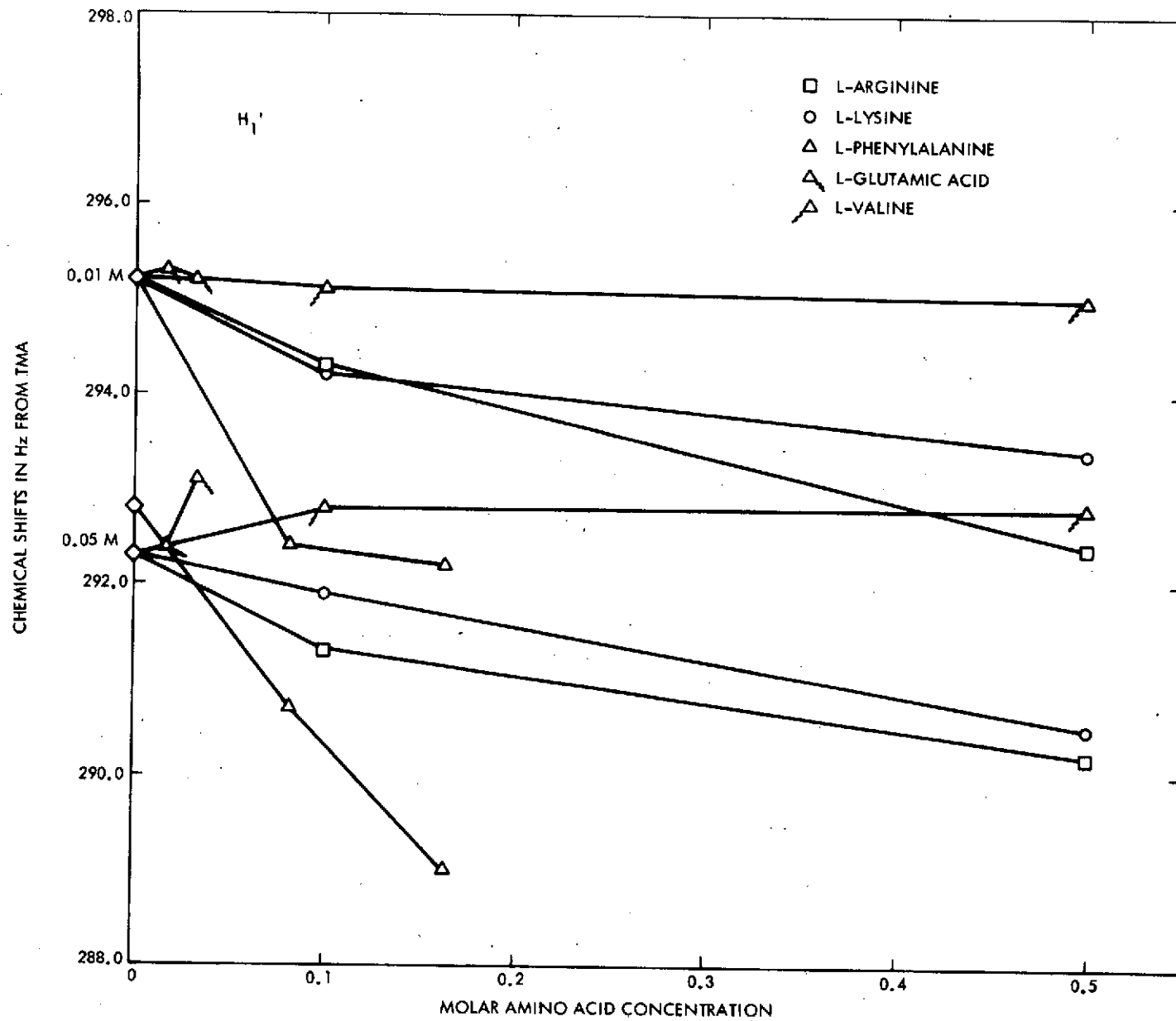


Fig. 11. Chemical shifts of  $H_1$ , of 5'-AMP in presence of amino acids,



which is somewhat surprising. This latter result suggests an association which withdraws electrons from the environment of  $H_8$  which is equally or slightly more important than the shielding due to intercalation of the aromatic ring of the amino acid.

In the case of glutamic acid, at the 5'-AMP concentration of 0.01 M, little effect is observed, while at the higher concentration (0.05 M), a slight deshielding arises for  $H_2$  and  $H_{1,}$ , but  $H_8$  is first slightly shielded and at higher amino acid concentration a definite deshielding takes place. Proton  $H_8$  is odd in that it exhibits first a significant shielding, then a deshielding.

At 0.01 M 5'-AMP, valine has essentially no effect on the  $H_2$ ,  $H_8$  and  $H_{1,}$  chemical shifts, while at 0.05 M 5'-AMP concentration it appears to cause a slight deshielding.

Both lysine and arginine appear to cause shielding for  $H_2$  and  $H_{1,}$ , at 0.01 M 5'-AMP.  $H_8$  is definitely deshielded. At the higher 5'-AMP concentration both these amino acids slightly shield  $H_{1,}$  and slightly shield  $H_8$  while no significant effect exists on  $H_2$  at 0.5 M lysine concentration and 0.5 M arginine slightly deshields  $H_2$ .

As mentioned above, the 5'-AMP is a polyfunction molecule which has several possible sites at which another molecule could interact. The difference in the observed chemical shift behavior of adenine and 5'-AMP protons with added amino acids reflect the difference in binding complexity. Besides the multiplicity of 5'-AMP binding sites, the possibility that the amino acids undergo significant dimerization and/or higher association in solution may need to be considered. There appears to be no data bearing on this possibility. In spite of the complex 5'-AMP behavior, the

stacking interaction between aromatic rings is evident. A better understanding of the complex situation with 5'-AMP and amino acids will require more of the type of data obtained in the present work and the following types of studies:

1. More results with adenine where potential contribution from amino acid self-interactions are considered.
2. Studies on adenosine with added amino acid.
3. Studies on adenine, adenosine and 5'-AMP where the amino end and then the carboxylate end of the amino acids are covered.
4. Studies on adenosine derivatives where the hydroxyl groups are masked with methyl groups.
5. Studies on 5'-AMP where the phosphate oxygens are masked with methyl groups.
6. Appropriate temperature studies.

Although there have been studies of the interaction of amino acids, their derivatives and small peptides with polynucleic acids there are only two reports of the studies involving interactions between nucleic acid constituents and polyamino acids.<sup>23,25</sup> Both of these studies suggested that specific interactions exist. As a last part of the preliminary proton NMR investigations undertaken in this work, we have studied the spectra of mixtures of three water soluble polypeptides and either 5'-AMP or adenine. We expected to observe chemical shift changes and/or line width changes which would signal the degree and some aspects of the type of any interactions existing. In the other results discussed above it should be

noted that no line broadening was observed in any case. Thus, on an NMR time scale the association rates between solutes in all these latter solutions are fast and any complex formed tumbles fast. In the case of proton NMR spectral studies of mononucleic acid constituent-polynucleic or oligonucleic acid interactions a number of cases of slower tumbling rates have been observed.<sup>50</sup> In no cases in the present work was line broadening

---

50. See for example: G. P. Kreishman and S. I. Chan, Biopoly., 10, 159 (1971) and ref. cited therein.

---

of the adenine or 5'-AMP or  $\alpha$ -C-H and  $\omega$ -CH<sub>2</sub>-amino acid proton signals noted. Thus, the fast exchange and tumbling situations seem to prevail here too. The particular polypeptides used in this work were chosen to be small. They were poly-L-proline degree of polymerization (DP) of 69, poly-L-lysine · HBr of DP 34 and poly-L-arginine · HCl of DP 72. It was hoped that small molecular weight polypeptides would give rise to sharper spectral lines and allow spectral changes to be sought more readily for the amino acid component. This indeed turned out to be the case.

Tables VI and VII summarize the NMR spectral data for the mixtures of the polypeptides and adenine and 5'-AMP. Figs. 12 and 13 allow comparison of the spectral changes of adenine and 5'-AMP as has been done for the monomer-monomer experiments discussed above. In the case of adenine, polylysine appears to cause little chemical shift change but, however, it definitely causes the resonance separation of the H<sub>2</sub> and H<sub>8</sub> protons to decrease. Both polyproline and polyarginine cause a definite shielding and also definite decrease in the spacings between the H<sub>2</sub> and H<sub>8</sub> proton

TABLE VI - Spectral Data

For Adenine And 5'-AMP Admixed With Polypeptides<sup>a</sup>

	Poly-L-proline			Poly-L-lysine		Poly-L-arginine	
	Molar Monomer Unit Concentrations For Polymers						
Adenine (0.0040 M)	0.0	0.022	0.085	0.024	0.099	0.023	0.095
H <sub>2</sub>	501.8 <sup>b</sup>	499.4	500.1	501.7	501.7	499.7	500.0
H <sub>8</sub>	497.5 <sup>b</sup>	495.6	496.1	497.5	497.9	496.0	496.7
(H <sub>2</sub> -H <sub>8</sub> )	4.3	3.8	3.9	4.2	3.6	3.7	3.3
5'-AMP (0.01 M)							
H <sub>2</sub>	536.5	536.8	535.9	533.2	534.7	528.4	529.1 (509.6) <sup>c</sup>
H <sub>8</sub>	506.5	505.7	504.7	500.1	503.6	497.8	498.7
(H <sub>2</sub> -H <sub>8</sub> )	30.0	31.1	31.2	33.1	31.1	30.6	30.4
H <sub>1'</sub>	295.5	294.7	292.7	291.5	293.0	289.7	286.6 (282.7) <sup>c</sup>
J <sub>12</sub>	6.1	5.9	5.8	5.8	5.8	5.6	5.6 (6.1) <sup>c</sup>

- a) Solutions prepared from half saturated solution of adenine at pH (observed) 7.08 and 0.02 M 5'-AMP at pH (observed) 7.05. Polypeptide stock solutions were adjusted to within  $\pm$  0.05 pH units of 7 and then mixed with an equal volume of either the adenine or 5'-AMP solution.
- b) All data for adenine taken at an 8° lower probe temperature than data in Table IV.
- c) There were some extra peaks present which indicated that the 5'-AMP had lost phosphate; the chemical shifts and coupling constant observed are appropriate for adenosine.

TABLE VII - Spectral Changes Of Protons Geminal  
To Nitrogen For Polypeptides

<u>Mixture</u>	$\delta_{\alpha\text{-CH}} - \delta_{\text{TMA}}^{a,b}$	$\delta_{\omega\text{-CH}_2} - \delta_{\text{TMA}}$	<u>Shift Due to Adenine or 5'-AMP<sup>d</sup></u>	
Polylysine 0.095 M	116.5	-15.7		
Polylysine 0.099 M and 5'-AMP 0.01 M	115.2	-17.0	+1.3	+1.3
Polylysine 0.024 M	108 <sup>c</sup>	-16.2		
Polylysine 0.024 M and 5'-AMP 0.01 M	105 <sup>c</sup>	-18.8	+3	+2.6
Polylysine 0.099 M and Adenine 0.0040 M	116.0	-15.2		+0.1
Polylysine 0.024 M and Adenine 0.0040 M		-16.1		-0.1
Polyarginine 0.091 M	118.5	+4.0		
Polyarginine 0.095 M and 5'-AMP 0.01 M	116.0	+1.2	+2.5	+3.8
Polyarginine 0.095 M and Adenine 0.0040 M	118.0		+0.5	
Polyarginine 0.023	119.0	+3.5		
Polyarginine 0.023 and 5'-AMP 0.01 M	116.5	-8.0	+2.5	+13.5
Polyarginine 0.023 and Adenine 0.0040 M	118.5	+3.5	-0.5	+0.0

- a) Plus sign means downfield from TMA.
- b) Could only be estimated to  $\pm 0.5$  Hz because of broadness of lines, overlap with sugar resonance and/or closeness to water line.
- c) Only estimated to  $\pm 1$  Hz.
- d) Plus sign means increase in shielding.



Fig. 12. Adenine chemical shifts changes in presence of polypeptides.

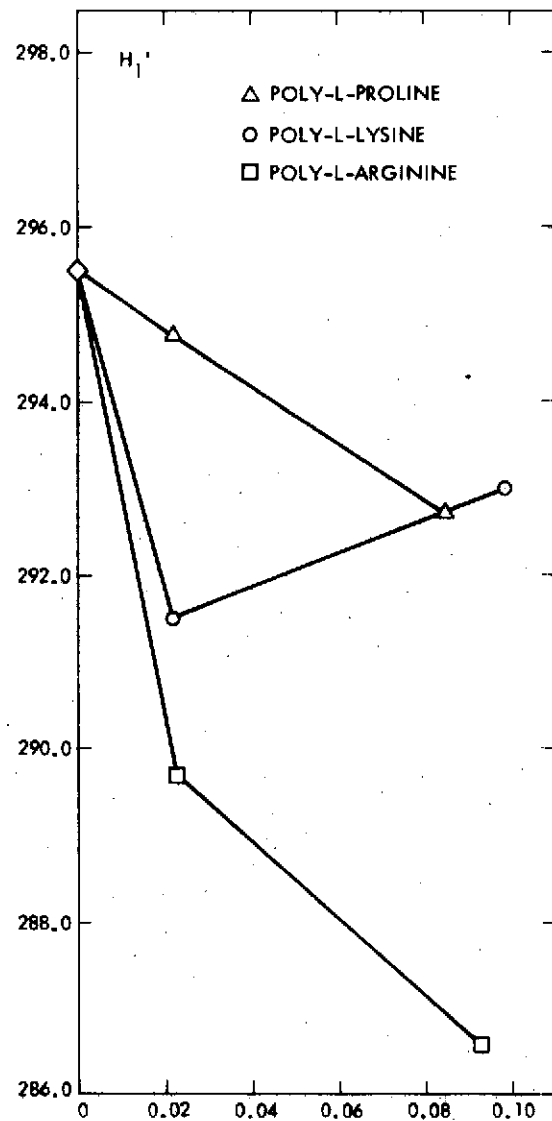
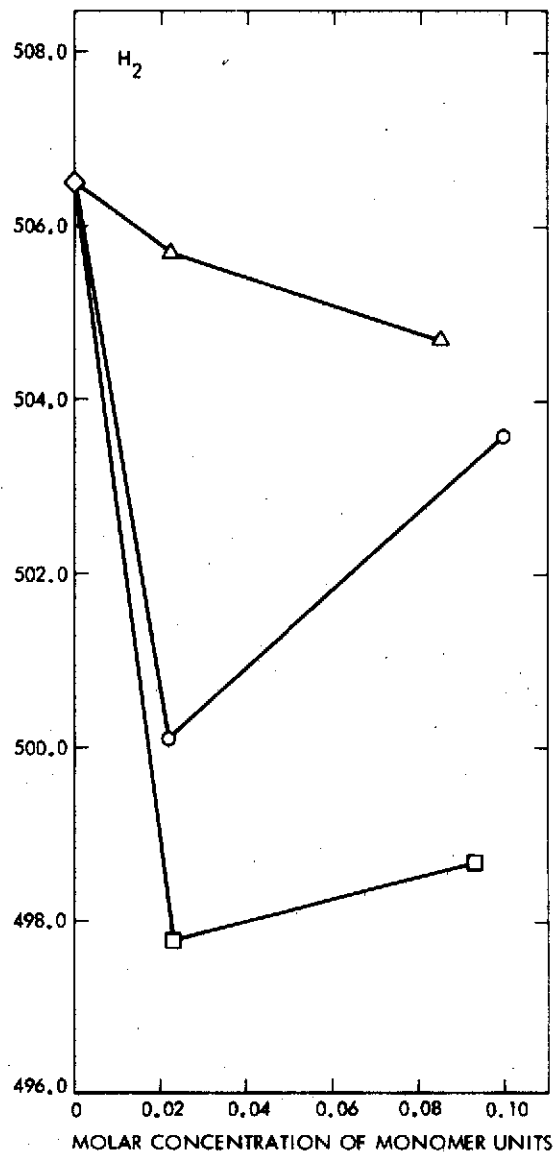
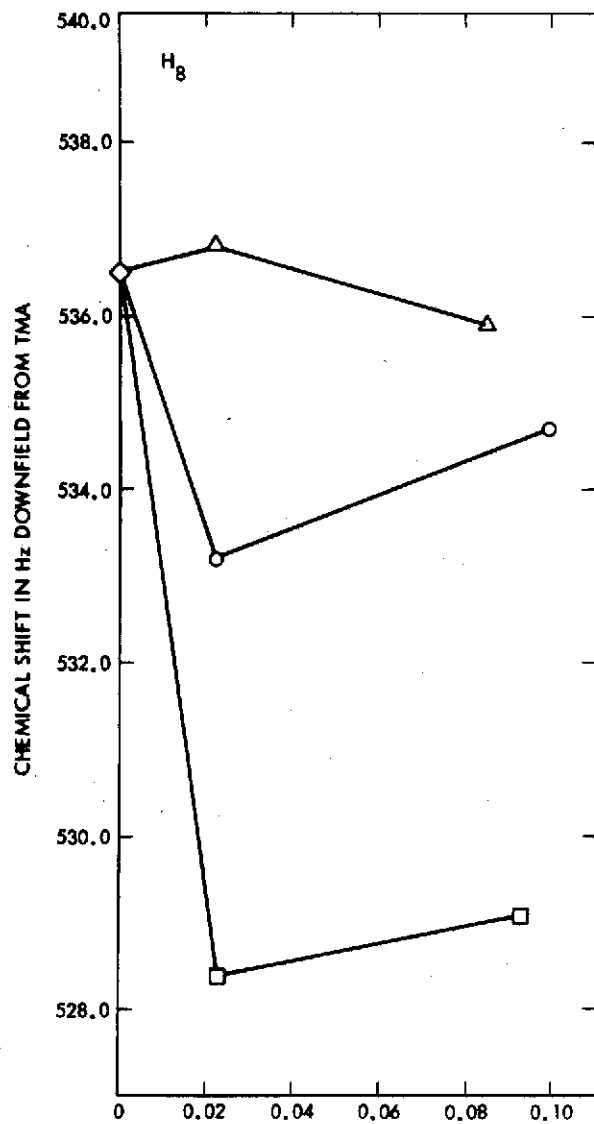
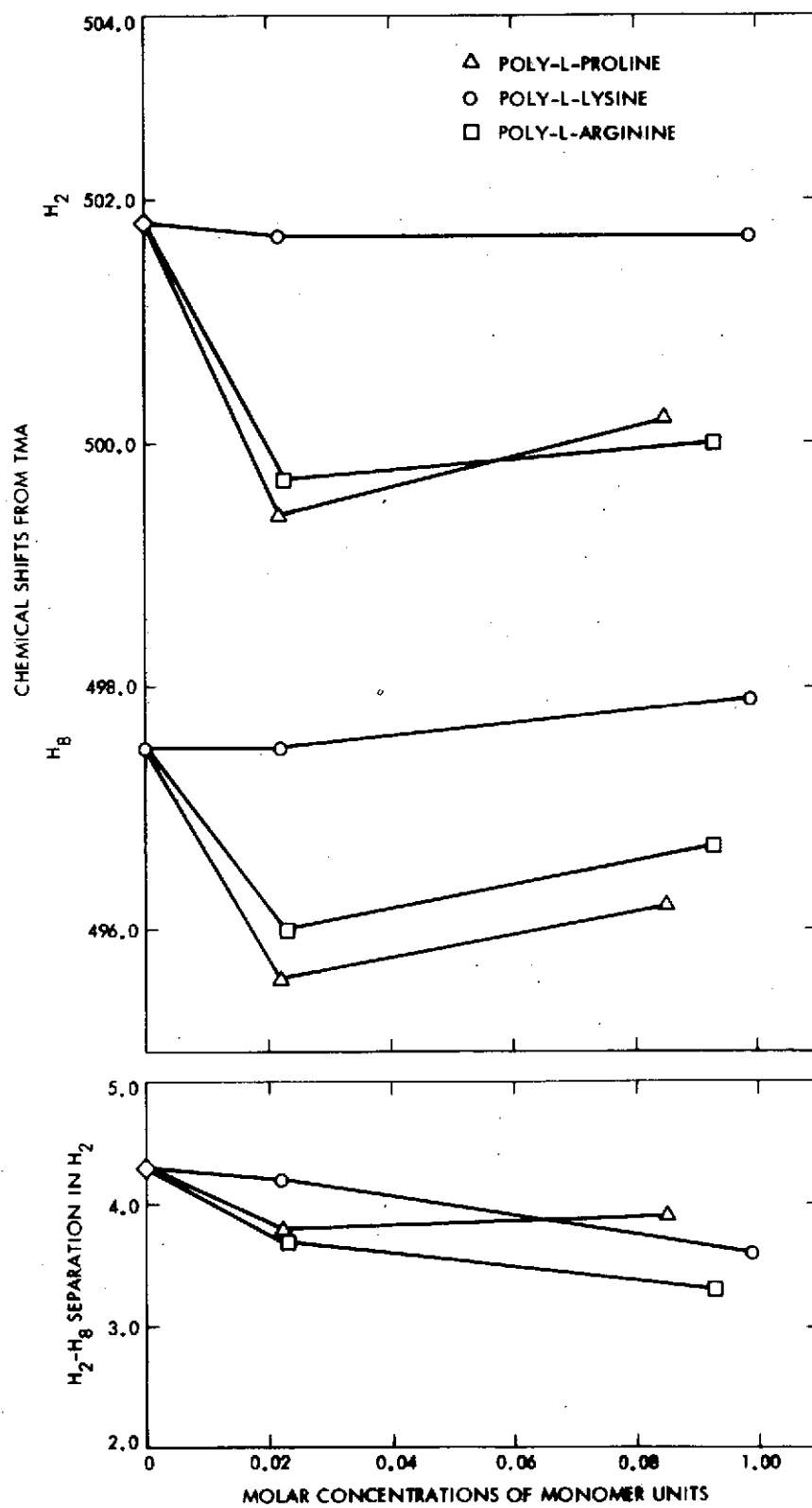


Fig. 13. Chemical shift changes of 5'-AMP  $H_8$ ,  $H_2$  and  $H_1$ , proton in presence of polypeptides.



signals. All three protons of 5'-AMP are affected similarly by polylysine. At lower concentration, a shielding increase is observed which becomes less of an increase at higher polymer concentration. In contrast polyproline has little effect on  $H_2$  but  $H_3$  and  $H_1$ , appear to undergo a nearly linear increase of shielding with concentration. Polyarginine increases the shielding of the 5'-AMP protons most of all. The behavior of  $H_2$  and  $H_3$  with this polymer is similar but of a lesser degree to that caused by polyproline. The  $H_1$  resonance is the most affected by polyarginine. It was noted as shown in Table VII that the  $\alpha$ -CH-N and  $\omega$ -CH<sub>2</sub>-N protons of polylysine and polyarginine exhibit increases in shielding in the presence of 5'-AMP. It is noteworthy that for a given 5'-AMP concentrations the largest increase is noted for the lower polypeptide concentration. In the case of polyarginine the  $\omega$ -CH<sub>2</sub>-N proton experience the largest increase of chemical shifts implying that the site of complexing might be closer to this proton. In Table VII the measurements of the  $\alpha$ -CH-N proton chemical shift positions are somewhat tentative because of signal-to-noise problems, this proton's closeness to the strong water peak and, because for the mixtures with 5'-AMP, certain sugar resonance signals fall in this same region. It was not possible to observe the  $\alpha$ -CH-N proton of polyproline mixtures because this signal appears to fall under the strong water signal. The N-CH<sub>2</sub> proton signals of this polymer give a broad band (line width 40 Hz) centered about 53 Hz downfield from the TMA internal reference. This broad band did not shift or change in line width on admixture with 5'-AMP.

<sup>14</sup>N NMR Studies of Several Amino Acids

Nitrogen NMR spectra can yield important information about the bonding structure and interactions of biologically important molecules. At 4.33 MHz the range of <sup>14</sup>N chemical shifts for amine nitrogen is comparable to the linewidth caused by quadrupolar relaxation. However, much information is contained in the relaxation time relative to molecular motion, chemical exchange and molecular environment. The primary objective of this part of the present work was to ascertain the possibility of utilizing <sup>14</sup>N relaxation times to obtain information regarding amino acid- nucleotide interactions. The amino acid spectra were studied as a function of pH to determine under what conditions any weak interactions might be studied. Our results suggested that the most fruitful studies should be done near the isoelectric point of the amino acid. Before detailed discussion of these results, we discuss some aspects of <sup>14</sup>N NMR spectra.

Nuclei with spin greater than 1/2 have their NMR spectra broadened by interaction of the nuclear quadrupole moment with a fluctuating electric field gradient at the nucleus due to molecular motion. For many purposes this broadening obscures information which may otherwise be contained in the NMR spectra. However, the relaxation time of a nucleus with spin greater than 1/2 can be a very sensitive function of the molecular environment. Stengle and Baldeschwieler<sup>42,51</sup> have used chlorine NMR to probe the

---

51. T. R. Stengle and J. D. Baldeschwieler, J. Am. Chem. Soc., 89, 3045 (1967).

---

reactivity of the sulfhydryl groups of hemoglobin and its fragments. Their method is based on the fact that the Cl linewidth,  $\Delta\nu$ , is proportional to the product of the correlation time for molecular reorientation, associated with quadrupole relaxation  $\tau_q$ , and the square of the Cl quadrupole coupling constant,  $e^2qQ$ ,

$$\Delta\nu \propto (e^2qQ)^2 \tau_q. \quad (14)$$

For a  $\text{Cl}^-$  ion, both terms in the product may change drastically when the  $\text{Cl}^-$  becomes bound to a large molecule. The change in linewidth for  $\text{Cl}^-$  is striking even when small amounts of  $\text{Cl}^-$  become bound, as the bound  $\text{Cl}^-$  may have a linewidth of many kHz. Only a small portion need be bound to cause an overall increase in linewidth of several Hz to several tens of Hz. The linewidth of the composite signal is given by

$$\Delta\nu = \Delta\nu_f P_f + \Delta\nu_b P_b \quad (15)$$

where the P's are the probabilities that the  $\text{Cl}^-$  is free or bound.

Recently, there have been some  $^{14}\text{N}$  NMR studies which indicate that the  $^{14}\text{N}$  nuclei in amino acids may be a satisfactory probe for investigating interactions between amino acids and polynucleotides, nucleosides, and nucleotides in solution. The  $^{14}\text{N}$  work has been of two types. In the first of these, one assumes a model for molecular motion and uses NMR data to derive the nitrogen nuclear quadrupole coupling constant.<sup>52</sup>

---

52. See for example: G. J. Jenks, J. Chem. Phys., 54, 1558 (1971).

---

The second approach is the inverse procedure where NMR and quadrupole data are used to infer a satisfactory model for molecular motion.<sup>53</sup> There are

---

53. A. Leowenstein and R. Waiman, Mole. Phys., 25, 49 (1973).

---

in the literature several discussion of the models which may be used to interpret <sup>14</sup>N linewidths. Ref. 53 cites the pertinent publications.

We will now consider a possible approach applicable to amino acid-nucleotide interactions. The spin-lattice relaxation time  $T_1$  for a <sup>14</sup>N nucleus (spin unity) is

$$T_1^{-1} = \frac{3}{8} \left(1 + \frac{1}{3} \eta^2\right) \left(\frac{e^2 q Q}{\hbar}\right)^2 \tau_q \approx T_2^{-1} \quad (16)$$

where  $\eta$  is the asymmetry parameter for the field gradient. The other terms have been defined previously. The correlation time for quadrupolar relaxation takes the form<sup>52,53</sup>

$$\tau_q = A \left(\frac{\sigma}{kT}\right) \quad (17)$$

where  $\sigma$  is the viscosity and A has the dimensions of a molecular volume.

There seem to be certain difficulties inherent in the use of amino acid <sup>14</sup>N nuclei as probes for interactions in solution. It would appear that it is necessary to know  $e^2 Qq$  for the <sup>14</sup>N nuclei in the solution under investigation in order to infer the correlation time. In fact  $e^2 Qq$  is not known in solution and is likely to be a strong function of pH. If however, we are interested primarily in changes in  $T_1^{-1}$  due to interaction of the amino acid with a polynucleotide or other complexing species, it is not necessary to have precise knowledge of  $e^2 Qq$  or  $\tau_q$ . If the amino acid



becomes associated with another molecule in solution and has its motion restricted within the complex, it seems likely that the correlation time for quadrupolar relaxation will increase by an amount related to the effective molecular volume of the complex. If, however, the amino acid retains some internal degrees of freedom the effect on  $\tau_q$  will be less. The quadrupole coupling constant for amino nitrogen is generally in the range of 3-5 MHz.<sup>52</sup> This will decrease at low pH<sup>54</sup> and when the nitrogen is in a tetrahedral

---

54. D. T. Edmonds and C. P. Summers, J. Mag. Res., 12, 134 (1973).

---

environment. It is likely on the basis of available data that  $e^2Qq$  will be in the range of 1-1.5 MHz in a tetrahedral environment.<sup>54</sup> A large range of values for  $\eta$  in solid amino acids occur.<sup>54</sup> However, it appears that the contribution to  $\eta$  from the crystal site will be lost in solution, so that a smaller range of values should exist for similarly solvated molecules. For fast exchange, it should be zero.

The <sup>14</sup>N linewidths for the amino acids and sec-butylamine (SBA) showed a striking pH dependence. The data are presented in Table VIII and plotted in Fig. 14. Similar behavior has been observed for glycine<sup>55</sup>

---

55. A. Tzalmona and E. Loewenthal, private communication; to be published.

---

and for histamine.<sup>44</sup> Tzalmona has explained the glycine results in terms of the protonation kinetics described by Sheinblatt and Gutowsky.<sup>56</sup> A

---

56. M. Sheinblatt and H. S. Gutowsky, J. Amer. Chem. Soc., 86, 4814 (1964).

---

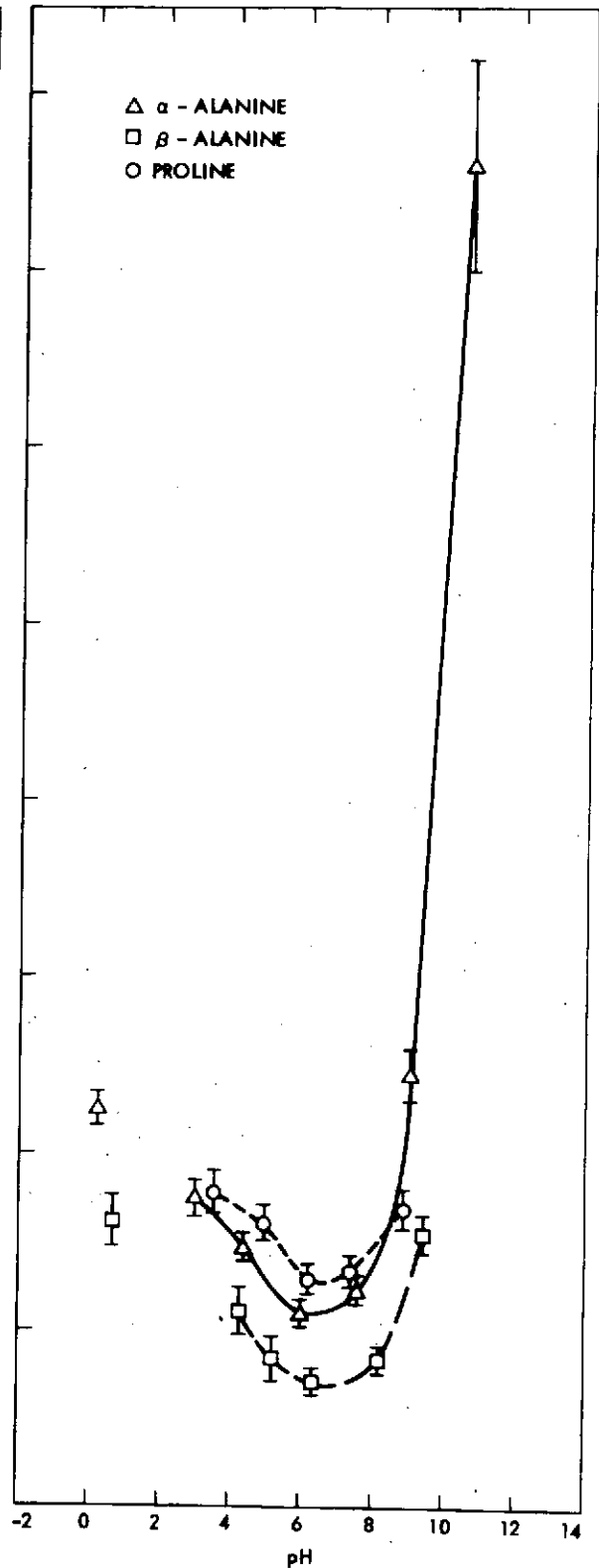
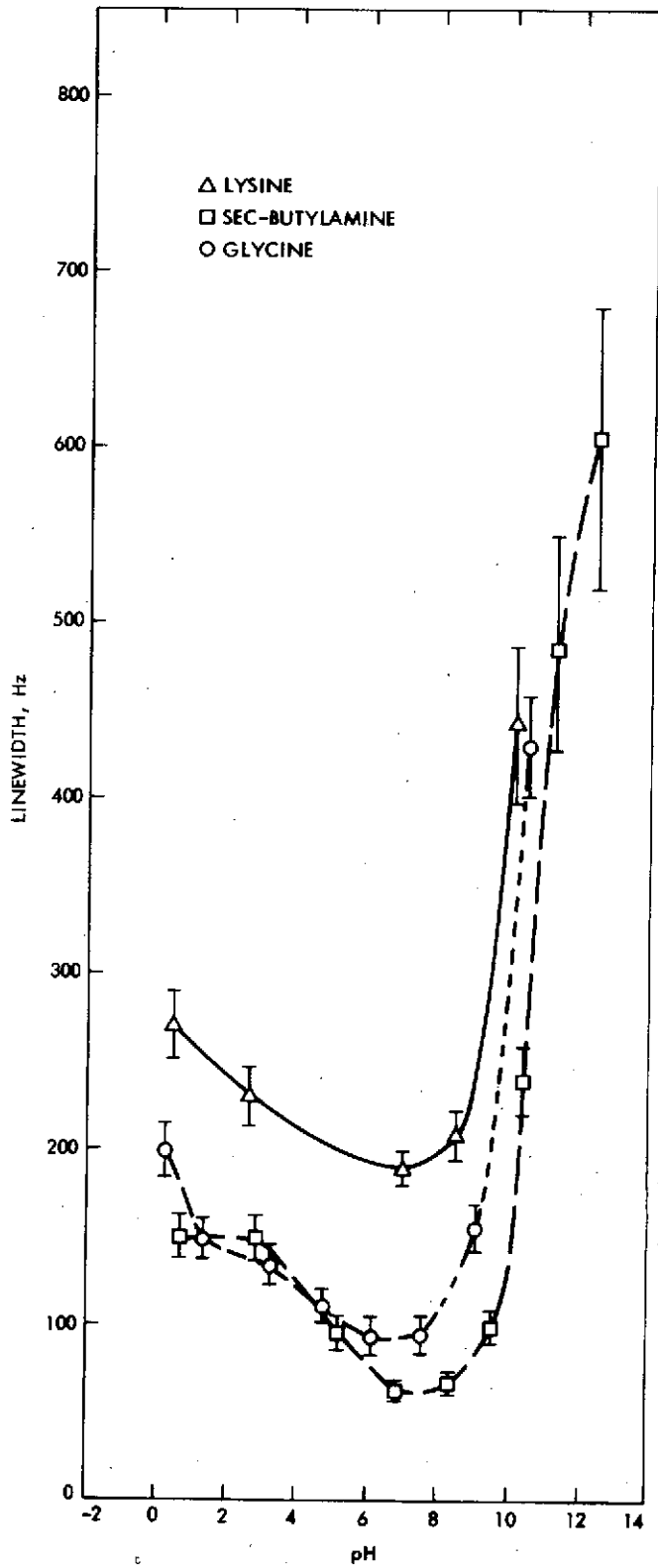
TABLE VIII -  $^{14}\text{N}$  Linewidth versus pH for Several  
Amino Acids and sec-butylamine

Glycine		$\alpha$ -Alanine		$\beta$ -Alanine	
<u>pH</u>	<u><math>\Delta\nu</math> (Hz)</u>	<u>pH</u>	<u><math>\Delta\nu</math> (Hz)</u>	<u>pH</u>	<u><math>\Delta\nu</math> (Hz)</u>
.27	199	.32	225	.03	162
1.35	152				
3.22	135	2.97	175		
4.76	113	4.35	148	4.28	112
5.93	94	6.05	111	5.40	86
7.62	96	7.54	124	6.36	72
9.07	153	9.20	246	8.14	84
10.51	430	10.05	760	9.45	156
Lysine		Proline		<u>sec</u> -butyl amine	
<u>pH</u>	<u><math>\Delta\nu</math> (Hz)</u>	<u>pH</u>	<u><math>\Delta\nu</math> (Hz)</u>	<u>pH</u>	<u><math>\Delta\nu</math> (Hz)</u>
.50	270	3.50	178	.69	148
2.67	230	4.89	160	1.30	150
7.06	190	6.20	130	2.87	148
8.53	207	7.23	135	5.20	96
10.13	440	8.84	170	6.86	75
				8.36	78
				9.58	100
				10.38	485
				12.44	606

Fig. 14. Amino acid  $^{14}\text{N}$  NMR linewidth-pH profiles.

- a. glycine, lysine, sec-butyl amine
- b.  $\alpha$ -alanine,  $\beta$ -alanine, proline

The shapes of the missing portions of the alanine curves have not been accurately determined.



recent study of the pH dependence of the  $^{15}\text{N}$ -glycine NMR spectrum at 6 MHz<sup>57</sup>

---

57. R. A. Cooper, R. L. Lichter and J. D. Roberts, J. Amer. Chem. Soc., 95, 3724 (1973).

---

is also consistent with the protonation kinetics discussed by Sheinblatt and Gutowsky. Roberts, et al.<sup>57</sup> also showed that the chemical shifts changes with pH will be of the order of several Hz in the acid region and 54 Hz between pH 6.6 and 13.6 (at 4.33 MHz). This is small compared to quadrupolar relaxation. The linewidth for  $^{14}\text{N}$  may be followed for the entire pH region. However, the  $^{15}\text{N}$  spectrum vanishes due to site exchange between pH 8.3 and 13.2. The  $^{14}\text{N}$  spectrum thus can be observed in some regions where the  $^{15}\text{N}$  spectrum cannot. The general shapes of all the pH-linewidth profiles observed in this work are consistent with what has been observed in the few cases studied before.<sup>55-57</sup> Our results are plotted in Fig. 14. The high pH portion can be explained by the equilibrium between the zwitter ion and  $\text{RNH}_2$ . A factor of 3-4 change in quadrupole coupling accounts for the order of magnitude increase in linewidth relative to that near pH 7. The acid region is dominated by exchange modulation of the NH coupling along with protonation of the carboxylate group.

There are several points to be made after inspection of Fig. 14. First, there is a reasonable range of pH values which results in moderately narrow  $^{14}\text{N}$  lines. This is in all cases near the isoelectric point of the amino acid and in the range of values which is realistic for biological systems. However, prebiotic condition could be more extreme. It is

reasonable to expect that amino acid-nucleotide base interactions can be reflected in the linewidth of the  $^{14}\text{N}$  amino acid resonance in this pH range. It had been one of the goals of this preliminary work actually to carry out such experiments. However, this proved impossible within the scope of the present work because of a failure of the magnet and power supply. However, the results obtained indicate that these experiments have potential for studying weak interaction. This is particularly true if, as some data indicate, the association constants<sup>32</sup> of the amino acids with the nucleotide bases in 1:1 complexes are of the order of  $0.1\text{-}10\text{ M}^{-1}$ . In this case, a large portion of the amino acid will be in equilibrium with a high molecular weight complex and should show an increase in linewidth. This may be an oversimplification, however. The amino acid is in multiple equilibria with the solvent and other amino acid molecules.<sup>56</sup> Curves such as those in Fig. 14 should be generated for a series of amino acid-nucleic acid constituent combinations.

As expected, the heavier amino acids have the broader lines. However, there is a striking difference in linewidth between  $\alpha$ - and  $\beta$ -alanine. This can be explained by a decrease in  $e^2qQ$  of about 20% going from the  $\alpha$ - to  $\beta$ -position with the amino nitrogen. This is not unreasonable and is not inconsistent with quadrupole coupling constants which have been estimated for the amino acids.<sup>58</sup> However, the experiment does not provide

---

58. Private communication from Professor Rex Richards.

---

a separation of this from changes in intramolecular interactions involving internal motion of the molecule.

EXPERIMENTAL SECTION

All substances used in this work were obtained commercially. The polypeptides were obtained from Miles-Yeda, Ltd: poly-L-lysine · HBr mol. wt. 7200, DP 32, lot #LY 170; poly-L-arginine · HCl mol. wt. 13,900, DP 72, lot #AR44; poly-L-proline, mol. wt. 6730, lot #PR17. Samples were prepared in distilled water or else D<sub>2</sub>O of about 99.5% or greater deuterium content which was obtained from several commercial sources (purity checked in most cases by NMR). Weighed amounts of solutes were mixed with known volumes of H<sub>2</sub>O or D<sub>2</sub>O in either 5 mm OD NMR tubes or in 2 or 5 ml volumetric flasks. For proton work, the D<sub>2</sub>O used contained a small amount of tetramethylammonium chloride (TMA) as an internal standard. Spectral line positions were measured relative to the resonance for the TMA ion because this species has been shown to experience much less interactions with purines in aqueous solution than the commonly used sodium 3-trimethylsilyl propane sulfonate.<sup>59</sup> The pH and pD of solutions prepared in volumetrics were

---

59. D. H. Live and S. I. Chan, Org. Mag. Res., 5, 275 (1973).

---

adjusted during mixing by adding very small volumes of concentrated base or acid (for D<sub>2</sub>O solution, NaOD and DCl) before the final estimated solution volumes were reached. It should be remembered that in the D<sub>2</sub>O solutions that pD = pH (meter reading) + 0.40.<sup>60</sup>

---

60. P. K. Gloscoe and F. A. Long, J. Phys. Chem., 64, 188 (1960).

---

Hydrated disodium ( $5 \cdot \text{H}_2\text{O}$ ) adenosine-5'-monophosphate (Sigma) was analyzed by UV and determined to be better than 95% pure. None of the samples were degassed; however, linewidths for the tetramethylammonium peak were less than 0.5 Hz.

Proton NMR spectra were recorded on a Varian HA-100 modified as detailed below. The "Sweep Oscillator Amplifier" and "Sweep Network" cards were removed from the V-4354A unit and in their place the output from a Hewlett-Packard 5100A synthesizer was fed into the "Sweep Oscillator Out" jack of this unit. Digital frequency sweep was obtained by use of a Barry Research LSC-7A linear sweep controller (available from Barry Research, 934 East Meadow Drive, Palo Alto, CA 94303) driven either from a 1 MHz or 2 MHz clock source. The digital programmer is configured to sweep a Hewlett-Packard 5100A frequency synthesizer over sweep widths of 10 to  $10^4$  Hz. The programmer sweeps the frequencies downward starting at a frequency one sweep width higher than the most significant of the digits to be swept. In the present proton work, the HDO signal in the samples was used for a control signal. Thus, the procedure was to lock on the HDO peak using the internal-lock system with a frequency of ca. 3000 Hz and to sweep an audio sideband. The audio-lock frequency was taken from a second HP-5100A synthesizer whose output is fed to the "Manual Oscillator Out" jack on the V-4354A; the "Manual Oscillator Tune Network" and "Manual Oscillator Amplifier" cards were removed from the V-4354A. In addition, the crystal of the V-4311 unit was removed and a 16.666666.66 MHz signal fed to this unit from a third HP-5100A synthesizer. For time averaging of spectra, an HP-5480 averager was utilized. This averager triggers the LSC-7A programmer whose



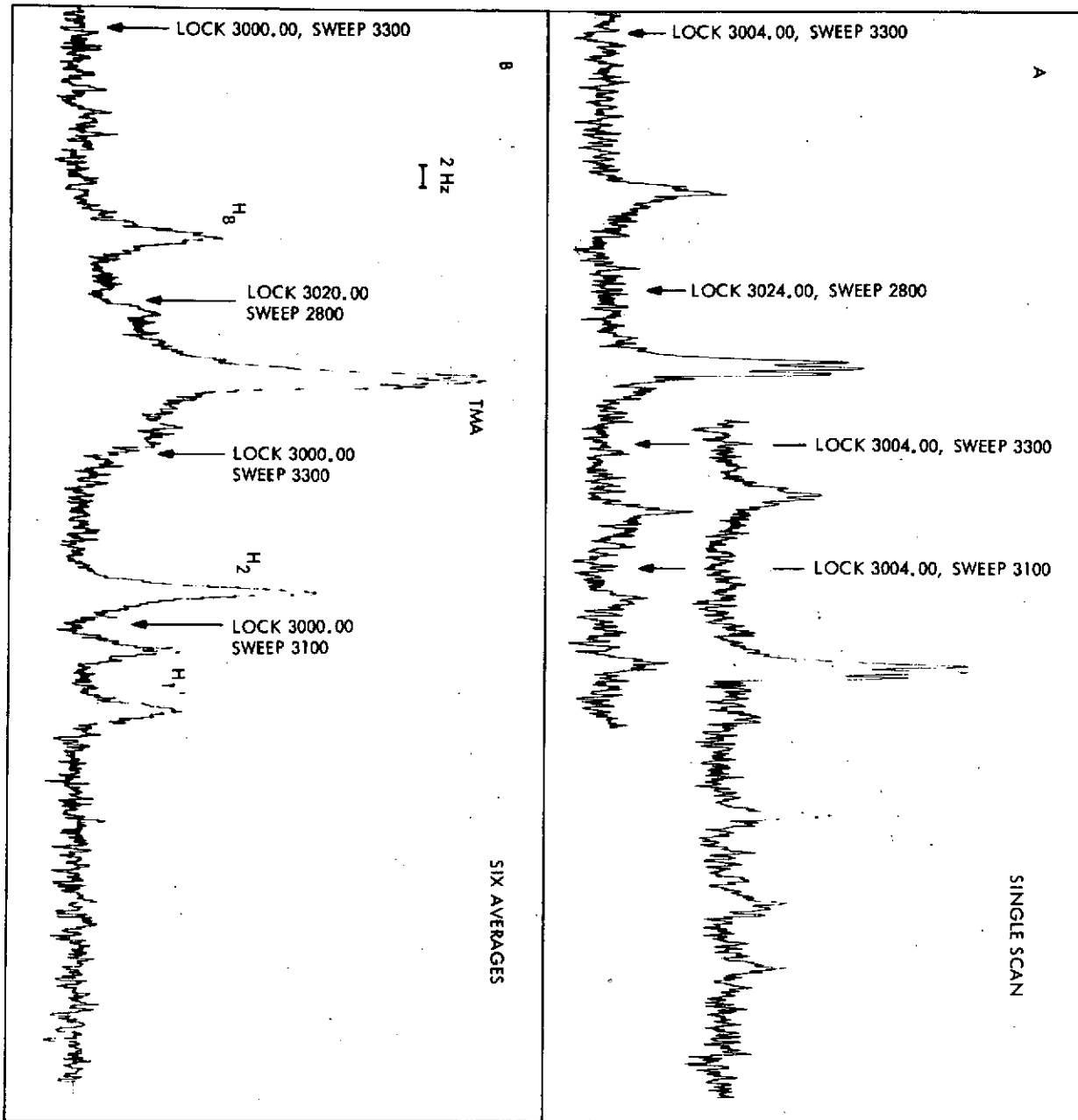
sweeptime was synchronized with the HP-5480 sweeptime. The accumulated spectra were either played out on the HA-100 chart recorder or else the spectra line positions could read off the display scope to  $\pm 0.1$  Hz.

The use of the digitally swept spectrometer described above enables highly accurate measurements of frequency differences between spectral lines many hundreds of Hz apart to be made with great facility. These measurements were made by the following procedure. The spectra lines for all the protons of interest in this work lie  $\pm 500$  Hz from the nominal 3000 Hz lock frequency. From rapid 500 Hz wide sweeps above and below 3000 Hz, the approximate ( $\pm 1$  Hz) position of all spectral lines in any sample were determined. In most cases, it was possible to record all the spectral lines of interest and the internal standard on one 100 Hz width scan. This was done by noting which 100 Hz width between 2500 and 3500 Hz any particular spectral line occurs and also its relative position in its 100 Hz sweep width. The programmer was started sweeping an appropriate 100 Hz width which would record the spectral line(s) lying at lowest field in a 100 Hz width. After this line(s) was recorded, the synthesizer button(s) appropriate to a 100 Hz width containing the next line or group of lines was (were) punched without interrupting the sweep. In cases where sets of lines in different 100 Hz widths would overlap, it was possible to instantaneously change the manual oscillator synthesizer frequency (and thus the spectrometer lock frequency) by exact increments so that the sets of lines could be recorded at different lock frequencies, but separated on a continuous 100 Hz frequency sweep. The same procedure was also used with the 5480 Averager.

Figure 15 exhibits such a spectrum showing all the spectral lines of interest for two 5'-AMP samples. Another measurement method used was to record in the averager separate 10 Hz or 100 Hz width spectra which contained spectral lines of interest at known lock frequencies. Although the spectrometer system described allows measurement of very precise spectral line positions, measurements in this present work were only made to  $\pm 0.2$  Hz because of signal-to-noise problems for many samples and the broadness of many lines. This accuracy was felt to be sufficient to exhibit spectral changes related to any weak interactions. It was recognized during this work that the position of the HDO signal used as a lock is very temperature sensitive, so that after insertion of a sample into the probe, spectra were not recorded for some minutes. Gross variations of room temperature resulting from a malfunctioning air conditioning system required that many of the time average spectra be remeasured. Thus, this air conditioning problem was probably the largest source of error in the spectral data reported herein, especially for related spectral measurements made hours to several days apart. It has been estimated that the largest error for related samples would be no more than  $\pm 1$  Hz. Probe temperature for most of this work was  $25 \pm 3^\circ$ .

The  $^{14}\text{N}$  studies were done on 1M solutions of glycine,  $\alpha$ -alanine,  $\beta$ -alanine, proline, lysine and sec-butylamine adjusted to various pH values between .3 and 12 with HCl and NaOH. These spectra were observed at 4.33 MHz using a Varian V-4311 RF unit and with either of two modulation schemes. For linewidths of less than 500 Hz, very low frequency, large amplitude squarewave modulation was used (5-10 Hz, 1-4.5 gauss). For linewidths

Fig. 15. Typical digital sweep 100 MHz proton spectra of two 0.01 M 5'-AMP solutions: A Two single scans with switches of lock frequency and sweep ranges as indicated; B Average of six scans of solution containing in addition 0.024 M poly-L-lysine. Both A and B are 100 Hz widths swept at a rate of 0.2 Hz/sec and one second time constant.



greater than 500 Hz, both the above scheme and a 50 Hz modulation frequency, low amplitude modulation (conventional wideline technique) were used with similar results. All spectra were frequency swept using a computer controlled (HP-2116B) HP-5100A Synthesizer. Averaging was accomplished synchronously in an HP-5480 Averager. No fieldlock was used as it was possible to maintain a drift small compared to the linewidth. A block diagram of the experimental setup is shown in Fig. 16. A typical spectrum is shown in Fig. 17. Linewidths were measured directly from the chart or the HP-5480 display. For large linewidths for which derivative shapes were obtained, computer integration and simulation were used to confirm the widths. Before any  $^{14}\text{N}$  studies could be accomplished with mixtures of amino acids and nucleic acid constituents, a catastrophic failure of one of the field coils and power supply of the spectrometer magnet took place; the duration and funding of present contract did not permit repair of this or reconfiguration of the rest of the spectrometer to another magnet.

The proton magnetic resonance work was performed by S. L. Manatt. The  $^{14}\text{N}$  spectrometer was assembled and the  $^{14}\text{N}$  studies of the amino acids were accomplished by E. A. Cohen. A majority of the careful preparation of the many solutions studies was done by A. M. Shiller. Helpful discussion with A. J. Bauman and P. A. Kroon contributed to this work.

Fig. 16. A block diagram of the 4.33 MHz  $^{14}\text{N}$  spectrometer

HP-2116 = computer  
HP-2759B = synthesizer programmer  
HP-5100A = synthesizer  
V-4311 = 4.33 MHz rf unit  
HP-211A = square wave generator  
PAR-HR8 = synchronous detector  
HP-5480 = signal averager

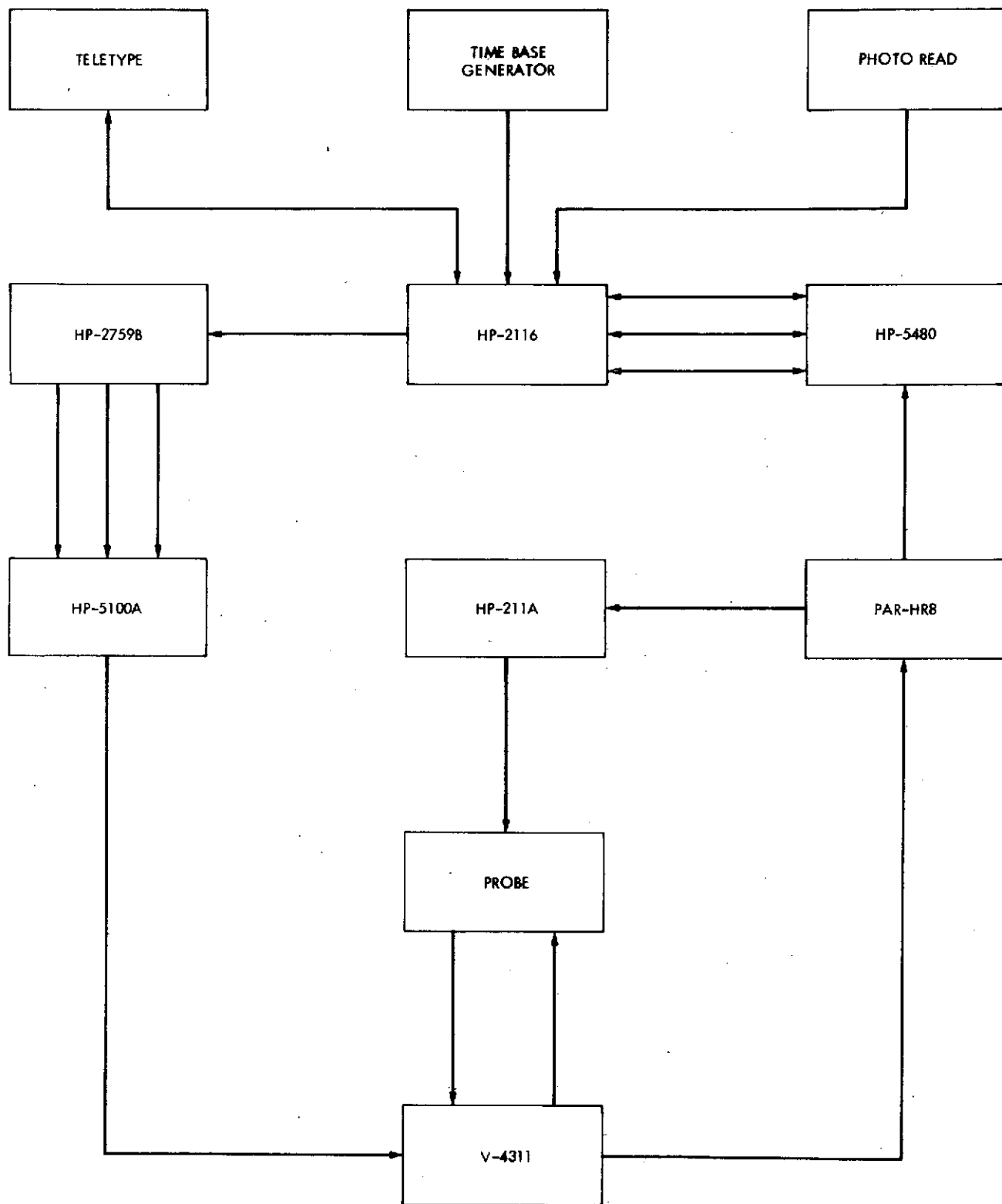
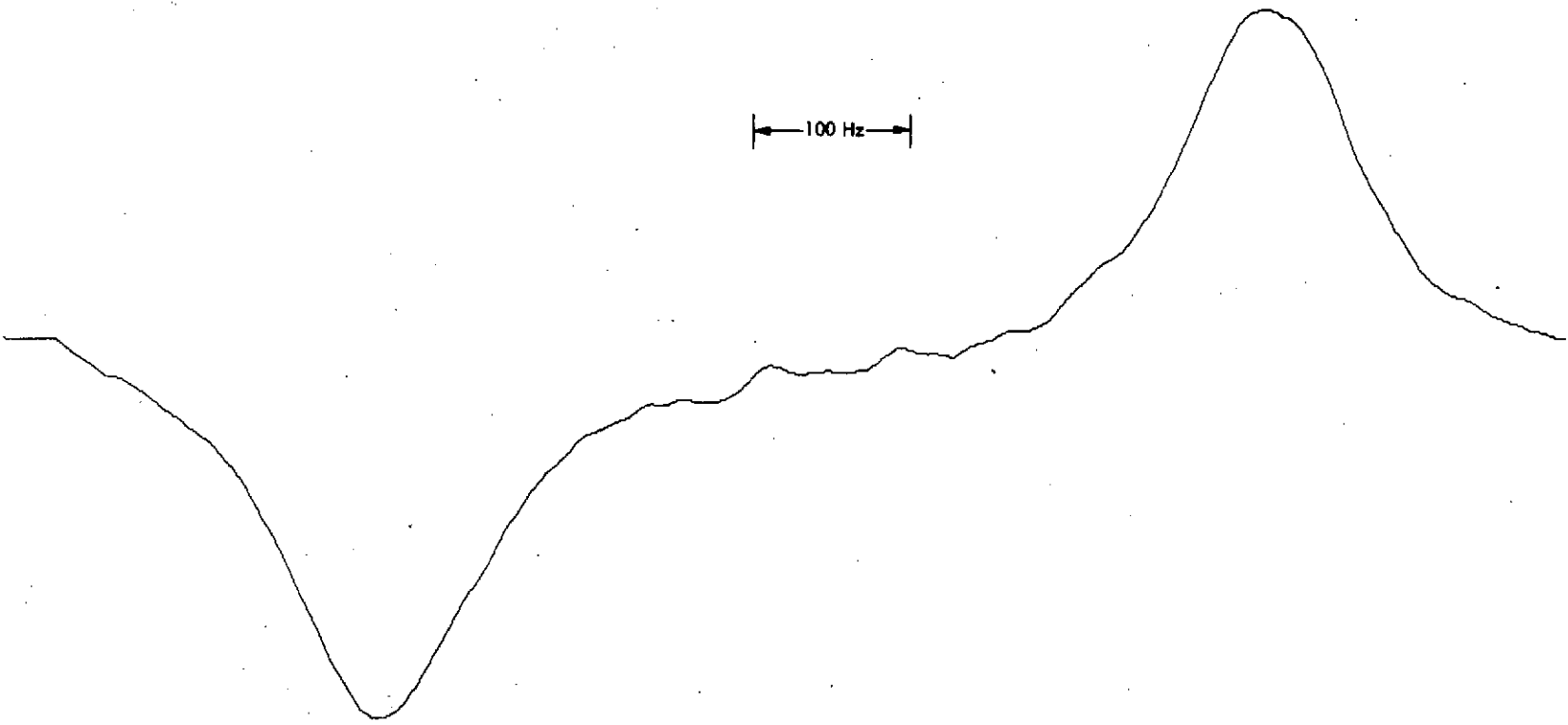


Fig. 17. The  $^{14}\text{N}$  NMR spectrum of  $\beta$ -alanine at pH 9.45.





CONCLUSION

The present work was undertaken to learn whether NMR techniques might successfully give information regarding the existence and extent of interaction between amino acids constituents and nucleic acid constituents in aqueous solution. The results show definitely that specific interactions do exist and can be studied by proton NMR. Although the number of results of the present preliminary study are limited, it is apparent that the quantitative extent of association between amino acid and nucleic acid components could be characterized in great detail by more extensive studies. The base stacking tendencies of aromatic rings of certain amino acids with nucleic acid constituents give rise to definite increased shielding of the proton signals of the nucleic acid component. Interaction of an acidic amino acid with cytosine and adenine deshields the corresponding protons. Tentatively, this has been attributed to a yet to be defined specific hydrogen bonding interaction. Neutral amino acids, lysine and arginine, seem only to deshield slightly. The observed effect could only be the result of slight destacking of the bases of the nucleic acid dimer or n-mer constituent or else the difference of two larger interactions of nearly equal and opposite shielding effects. Experiments to distinguish between these possibilities were not done. It was pointed out that studies with 5'-AMP and added amino acid should be rather complicated because the former has several potential sites for weak association. Definite chemical shift evidence of the effects of complex formation between three polypeptides and 5'-AMP and three polypeptides

C-2

and adenine was found. Thus, both  $\alpha$ -CH-N and  $\omega$ -CH<sub>2</sub>-N type protons are shielded more by added 5'-AMP. The behaviors of the H<sub>2</sub>, H<sub>8</sub> and H<sub>1</sub>, of the latter and H<sub>2</sub> and H<sub>8</sub> of adenine are rather complex.

In the case of cytosine-tryptophan, an association constant  $K = 2 \text{ M}^{-1}$  was estimated. An extended analysis of chemical shifts of the two adenine signals was presented from which it was possible to assign H<sub>2</sub> as the downfield proton signal. This conclusion was confirmed by partial deuterium exchange at H<sub>8</sub>. A lower limit for the dimerization constant of adenine was made ( $K = 8 \text{ M}^{-1}$ ). From the three sets of chemical shifts at three concentrations of 5'-AMP, the dimerization constants of the latter were estimated as  $K_2 \approx 2$ ,  $K_8 \approx 3$  and  $K_1 \approx 1 \text{ M}^{-1}$  with the following reasonable dimer chemical shifts of 40, 53, and 48 Hz, respectively.

A discussion of the potential of using <sup>14</sup>N NMR to study weak complexes between amino acids and nucleic acid constituents was presented. As a prerequisite to such studies, the pH dependence of the <sup>14</sup>N linewidth of several amino acids is reported for the first time. These results suggest that each amino acid has its own unique linewidth vs. pH profile which is probably determined by local electric field gradients and exchange effects. It was concluded that the linewidth should be significantly affected by weak association with added molecules such as nucleic acid derivatives.

The power of the NMR technique suggests a number of types of experiments to further understand interactions between amino acid and nucleic acid constituents. Many of these are mentioned above. Perhaps one of the

most interesting lines of work would be studies involving trinucleotides and some higher oligonucleotides with amino acids and small peptides to test some of the recent suggestions of Lacey,<sup>24a</sup> and Gabbay<sup>37,38</sup> regarding the significance of polypeptide-nucleotide interactions in evolution of the synthesis of biomolecules and the contemporary protein synthesis mechanism. However, first a much more detailed characterization of these interactions on the monomer level needs to be accomplished.