PROTEIN–NUCLEIC ACID INTERACTIONS

A THESIS

Submitted to the University of Roorkee
for the award of the degree
of
DOCTOR OF PHILOSOPHY

By

SHRIKANT KUKRETI

DEPARTMENT OF BIOSCIENCES & BIOTECHNOLOGY
UNIVERSITY OF ROORKEE
ROORKEE-247 667 (INDIA)

FEBRUARY, 1989

DEDICATED

TO MY FATHER

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled, **'PROTEIN-NUCLEIC ACID INTERACTIONS'** in fulfilment of the requirement for the award of the Degree of Doctor of Philosophy in Department of Biosciences and Biotechnology, University of Roorkee, Roorkee is an authentic record of my own work carried out during a period from September 1985 to January 1989, under the supervision of Dr. (Mrs.) Ritu Barthwal.

The matter embodied in this thesis has not been submitted by me for the award of any other Degree.

Shrikant Kukreti

(SHRIKANT KUKRETI)

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Ritu Barthwal

Dr. (Mrs.) RITU BARTHWAL)
Reader & Head
Department of Bioscience & Biotechnology
University of Roorkee
Roorkee 247 667, India

Head
Department of Biosciences
and Biotechnology
University of Roorkee
ROORKEE-247 667

Dated : 2.1.1989

The candidate has passed the Viva-Voce examination held on Oct 23, 1989 at Roorkee The thesis is recommended for the award of Ph.D. Degree.

Ritu Barthwal
Signature of Guide

23.10.89

Girjesh Govil
Signature of External Examiner(s)

ABSTRACT

Model systems involving nucleic acid and protein constituents have been widely used to obtain information on the specific interaction involved in protein-nucleic acid associations. The advantage to carry out this type of studies is the possibility to investigate interactions of specific nucleotide sequence of DNA with peptide chain with controlled amino acid composition, sequence, chain lengths and predictable conformation, which results in considerable simplification of the system as compared to the natural protein-DNA systems. Stacking of aromatic amino acids with nucleic acid bases is found to exist in the crystal structure of nucleoside peptide 5-[N-(L-phenylalanyl)-amino] uridine. Tyrosine, phenylalanine and tryptophan are found to stack in complexes of prokaryotic ribonuclease T with 2'-GMP, gene 5 protein of phase fd with (dA)5, lac repressor with operator etc. The present study has been undertaken to investigate stacking interactions using model systems with a view to find base/base sequence specificity and conformational changes induced on stacking. We have studied interaction of Tyr, Trp, Phe containing tri- and tetrapeptide with d-CpG d-GpC, d-GpCpGpC and d-CpCpGpG. One-dimensional NMR is used to find chemical shifts due to stacking, changes in Tm etc. while two-dimensional NMR techniques-J-correlated spectroscopy (COSY) and nuclear overhauser enhancement spectroscopy (NOESY) serve as methods to assign all proton NMR signals unambiguously and determine solution conformation such as sugar pucker,

helix sense, glycosidic bond rotation and interproton distances. We also present results of theoretical energy calculations made on stacking of Tyr, Trp, Phe with C, G bases, A-T and C-G base pairs and intercalation between C-G and G-C base pairs. Classical potential function has been used to estimate conformational energy. Total interaction energies are calculated as sum of electrostatic, dispersion, polarization and repulsion terms.

The upfield shifts of aromatic ring proton resonances of Trp, Tyr and Phe indicate stacking. The interaction is more prominent with double-helix than with single-stranded DNA. Practically in all cases DNA is double-helix with and without binding, sugar conformations are mostly O1'-endo, glycosidic bond rotations are anti/high anti. Interproton NOE's show most of the peptides are either in close proximity or stack with nucleotides. Among oligopeptides, the binding of d-GpCpGpC with Lys-Tyr-Lys is strongest. The binding of d-CpG with various aromatic amino acids of peptide have been found to be in the order Trp > Tyr > Phe. From theoretical studies, the interaction energies for model system d-CG with aromatic amino acids are in the same order and thus are in agreement with NMR results that we have obtained.

From our studies, we have thus shown that there is a specified type of base/base sequences specificity in nucleic acid-protein interactions. This may be the reason for specific recognition of binding site of nucleic acid by proteins.

ACKNOWLEDGEMENTS

I take this opportunity to express my sincerest sense of gratitude to my supervisor Dr. (Mrs) Ritu Barthwal, Reader and Head, Department of Biosciences and Biotechnology, University of Roorkee, for her erudite guidance, invaluable help and encouragement without which it would have been impossible for me to complete this work. It is a rich experience to have worked with her. I shall remain indebted to her forever.

My profound and sincere thanks are due to Prof.C.B.Sharma (Ex-Head of the Department) for providing necessary research facilities and constant encouragement during the course of research work.

I can not find adequate words to express my feelings for my friend and colleague Mr. Anwer Mujeeb Azmi who has been by my side all through, helping and encouraging me all the time. It would be rather difficult to find as sincere a colleague as he has been. I also acknowledge gratefully the help of Miss Anita Gupta during the preparation of this dissertation.

My tender sentiments are for all my affectionate friends-BADS - and colleagues in particular to Prasoon Chaturvedi, Rakesh Mohindra and Santosh K. Tripathi, who helped in one way or other.

My personal and deep appreciation goes to my colleague and friend Ritushree for her thoughtful co-operation during entire work span.

I shall ever be obliged by the help extended by the FT-NMR National Facility at T.I.F.R. Bombay for providing me the necessary and a friendly environment during my stay there.

Financial assistance rendered by Department of Science and Technology, Govt. of India, is gratefully acknowledged.

I am short at words in expressing my gratitude to my parents whose blessings, affection, cheerful and enthusiastic efforts, have been a constant source of inspiration. Love of my brothers Shashi and Swasti is gratefully acknowledged.

Shrikant Kukreti

(SHRIKANT KUKRETI)

CONTENTS

		Page No.
	CANDIDATE'S DECLARATION	ii
	ABSTRACT	iii
	ACKNOWLEDGEMENTS	v
Chapter - 1	INTRODUCTION	1
	- General	1
	- Structure of Nucleic Acids	2
	- Elements of Protein Structure	13
	- Protein-Nucleic Acid Complexes	15
	DNA Binding Proteins	15
	RNA Binding Proteins	20
	- Nature of Interaction Between Functional Groups In Protein-Nucleic Acid Associations	23
	1. Electrostatic Interactions	24
	2. Hydrogen Bonding Interactions	26
	3. Hydrophobic Interactions	31
	4. Stacking Interactions	33
	A. NMR studies	34
	B. Fluorescence studies	35
	C. Theoretical studies	37
	- Scope of Thesis	38
Chapter - 2	MATERIALS AND METHODS	41
	- Materials	41
	- Sample Preparation	41
	- NMR Methods	43
	A. Theory	43

		Page No.
	B. Experimental	53
	C. Strategies used for Conformational Analysis	54
	- Model Building	64
	- Theoretical methods	67
Chapter - 3	d-CpG AND ITS BINDING TO OLIGOPEPTIDES CONTAINING Trp, Tyr AND Phe RESIDUES	73
	- d-CpG	73
	- Binding of d-CpG to Lys-Trp-Gly-Lys OtBu	91
	- Binding of d-CpG to Lys-Tyr-Lys	105
	- Binding of d-CpG to Lys-Phe-Lys	117
Chapter - 4	d-GpCpGpG AND ITS BINDING TO TRIPEPTIDE Lys-Tyr-Lys	125
	- d-GpCpGpC	125
	A. Assignments	125
	B. Conformation	135
	- Binding of d-GpCpGpC to Lys-Tyr-Lys	140
Chapter - 5	d-GpC AND ITS BINDING TO Lys-Tyr-Lys AND Lys-Trp-Gly-Lys OtBu	157
	- d-GpC	157
	- Binding of d-GpC to Lys-Trp-Gly-Lys-OtBu	164
	- Binding of d-GpC to Lys-Tyr-Lys	175
Chapter - 6	d-CpCpGpG AND ITS BINDING TO TRIPEPTIDE Lys-Tyr-Lys	188
	- d-CpCpGpG	188
	A. Assignments	188

		Page No.
	B. Conformation	202
	- Binding of d-CpCpGpG to Lys-Tyr-Lys	203
Chapter - 7	STACKING ENERGY CALCULATIONS FOR Trp, Tyr, Phe	218
	- Charge Distribution On Interacting Molecules	218
	- Minimum Energy Conformation	218
	- Interaction With Bases	223
	- Interaction With Base Pairs	228
	- Interaction with C-G (Above) and G-C (Below) Base Pairs	231
	- Specific Recognition By Stacking Interaction	237
Chapter - 8	CONCLUSIONS	239
	REFERENCES	247

CHAPTER - I

INTRODUCTION

GENERAL

Proteins and nucleic acids, the two classes of biological molecules, play the key role in all cellular processes. The main function of nucleic acids is the storage and faultless transmission of genetic information to progeny and transmission and coding of this information into proteins whereas proteins are concerned with the execution of biological processes. The function of nucleic acids is controlled and made possible by their interactions with specific proteins. Thus the interactions between the proteins and nucleic acids play a very fundamental role in living cells, by controlling every step of the entire mechanism of information transfer including replication, transcription, translation, repressor mechanisms and presumably differentiation. In addition, binding of proteins to nucleic acids result in macroscopic structures like virus, ribosomes, nucleosomes etc.

Nucleic acid binding proteins which have no enzymatic activity may have a structural role by maintaining a given conformation or folding of the nucleic acid (e.g. nucleosomes) or they may interfere with enzymes involved in replication or transcription (e.g. single–strand binding proteins interacting with DNA polymerases, repressor or CAP protein interacting with RNA polymerase. Most often the protein recognises

a structure (e.g. single-strand binding proteins) or a nucleic acid base sequence (e.g. repressors). The process by which proteins specifically bind to well defined structures or base sequences is referred to as selective recognition. There may be a direct recognition of base sequence or an indirect mechanism involving cooperativity between proteins for binding to single-stranded structure. Sequence specific proteins always have some affinity for non specific sites. Therefore, selective recognition of the specific sequence may involve many preliminary steps during which the protein wanders about on non specific sites before being trapped on its specific site. When the nucleic acid binding protein exhibits an enzymatic activity the specificity of the overall reaction might not rest in the association step but in the catalytic steps which follow this association. In such cases the specificity will find its origin in the kinetics of the catalytic events rather than the selective recognition of the nucleic acid by the enzyme. An example is provided by the tRNA-aminoacyl-tRNA synthetase system. In case of restriction endonucleases both selective recognition and kinetic specificity most probably contribute to the specificity of the cleavage reaction for DNA sites of well defined sequences.

STRUCTURE OF NUCLEIC ACIDS

Nucleic Acids are formed by polymerisation of nucleotide units which contain a sugar phosphate backbone and one of

the four bases Adenine (A), Thymine (T), Guanine (G), Cytosine (C) in DNA and Adenine (A), Uracil (U), Guanine (G), Cytosine (C) in RNA [123, 51]. The sugar is either a ribose (RNA) or a deoxyribose (DNA). The base A pairs with T and G with C in double helical DNA in a geometry shown in figure 1.1. The sugar phosphate backbone unit is also shown in the figure. The backbone contains seven single bonds yet there is high degree of rotational hindrance primarily due to the repulsive interactions between negatively charged oxygen atoms and to the high correlation between the different rotations [138]. Consequently the backbone torsional angles $\omega, \Phi, \Psi, \Psi', \Phi', \omega'$ (other nomenclatures also exist in literature) in sugar phosphate backbone; the torsional angles, $\tau_0, \tau_1, \tau_2, \tau_3, \tau_4$ and hence the structure or the puckering of the five membered sugar ring, and the glycosyl torsional angle χ_{CN} which defines the orientation of the base relative to the ribose unit adopts specific preferred values and hence specific conformations (see Table 1.1 for definition of torsional angles).

The precise conformation of a molecule and can be defined in terms of the value of rotational angles around chemical bonds in question. For a four atoms segment A-B-C-D of a molecule, the torsional angle Ø for the bond B-C is defined as the angle between the normals to the planes A-B-C and B-C-D. A value of zero is assigned to the dihedral angle for a structure in which atom A and D are eclipsed. Clockwise rotation of C - D bond relative to A - B bond when viewed along

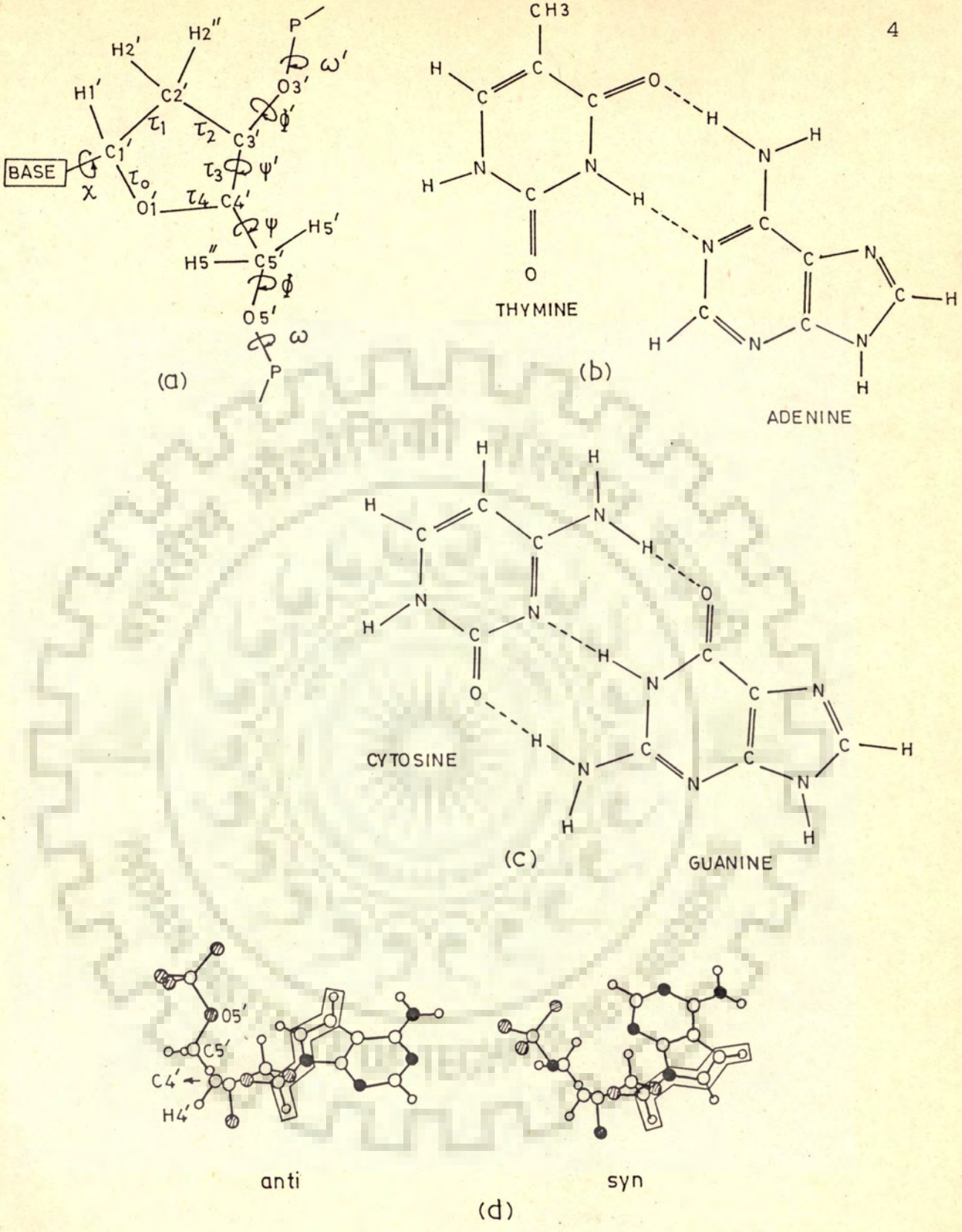


Fig. 1.1 **(a)** The nucleotide unit **(b)** A-T and **(c)** C-G base pair geometries and **(d)** Anti-syn orientations about the glycosyl bond.

Table 1.1a : Definition of Torsion Angles in Nucleotides [51][a]

Torsion angle	Atoms involved
ω	03'-P-05'-C5' (n-1)
Φ	P-05'-C5'-C4'
ψ	05'-C5'-C4'-C3'
ψ'	C5'-C4'-C3'-03'
Φ'	C4'-C3'-03'-P
ω'	C3'-03'-P-05' (n+1)
χ	01'-C1'-N1-C6 (Pyrimidines)
	01'-C1'-N9-C8 (Purines)
τ_0	C4'-01'-C1'-C2'
τ_1	01'-c1'-C2'-C3'
τ_2	C1'-C2'-C3'-C4'
τ_3	C2'-C3'-C4'--01'
τ_4	C3'-C4'-01'-C1'

[a]Atoms designated (n-1) and (n+1) belong to adjacent units.

Table 1.1b : Conversion of Different Definitions for Torsion Angle about the Glycosyl C1'-N Linkage [$\chi_{present} = \chi_{other} + Difference$] [123]

	Present	Other definition	Difference between present and other
Purine		04'-C1'-N9-C8	+180°
Pyrimidine		04'-C1'-N1-C6	
Purine	04'-C1'-N9-C8	C2'-C1'-N9-C8	-62.5°
Pyrimidine	04'-C1'-N1-C2	C2'-C1'-N1-C6	
Purine		C2'-C1'-N9-C4	+116.5°
Pyrimidine		C2'-C1'-N1-C2	

B - C axis is considered to be positive. The various conformational states are defined as

Conformation	Range of dihedral angle (in degrees)
Syn	0 ± 90°
anti	180 ± 90
cis	0 ± 30
trans (t)	180 ± 30
gauche$^+$ (g)	60 ± 30
gauche$^-$ (g-)	300 ± 30

The cis conformation refers to an eclipsed geometry whereas trans and anti are described as extended structures. On the other hand, gauche$^+$ and gauche$^-$ conformations refer to the staggered molecular geometries. The three dimensional structure of any molecule containing N atom can be defined in terms of N-1 bond length, N-2 bond angles and N-3 dihedral angles or alternately by 3N coordinates of atoms. Since the force constant for stretching and bending vibrations are relatively large the variations in the bond length and bond angles are not significant. Assuming standard bond lengths and bond angles for various bond types, the 3-dimensional structure of a molecule can be conveniently described by the N-3 dihedral angles.

The conformation of sugar moiety type plays a crucial role in determining the secondary and tertiary structure of nucleic acids. The steric factors make the furanose ring non planar. Being a cyclic structure the five torsional angles $\tau_0, \tau_1, \tau_2, \tau_3$ and τ_4 are highly correlated and any two can define the sugar geometry unambiguously. Variation of these dihedral angles result in an infinite number of conformations which are elegantly described in a unified way by applying the concept of pseudorotation [122, 51]. The torsional angles are described in terms of the phase angle (P) of pseudorotation and the degree of pucker (τ_m) as

$$\tau_j = \tau_m \text{ Cos}[P + (j-2)\delta]$$

where j takes values between 0 to 4 and τ_m and δ are 38° and 144° respectively. The phase angle (P) is related to the dihedral angles as

$$\tan P = \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2.\tau_3.\ (\text{Sin } 36° + \text{Sin } 72°)}$$

As P goes through a complete pseudorotation cycle, i.e. 0 to 360°, values of $\tau_0 - \tau_4$ range from $-\tau_m$ to $+\tau_m$ through 0° (see fig.1.2). Some of the classical sugar ring puckers are represented on the pseudorotation wheel in Fig.1.2. Analysis of crystal structure data show that two major sugar puckerings i.e. C3'-endo and C2'-endo are commonly found which correspond to P = 18° to 162° respectively in the pseudorotational

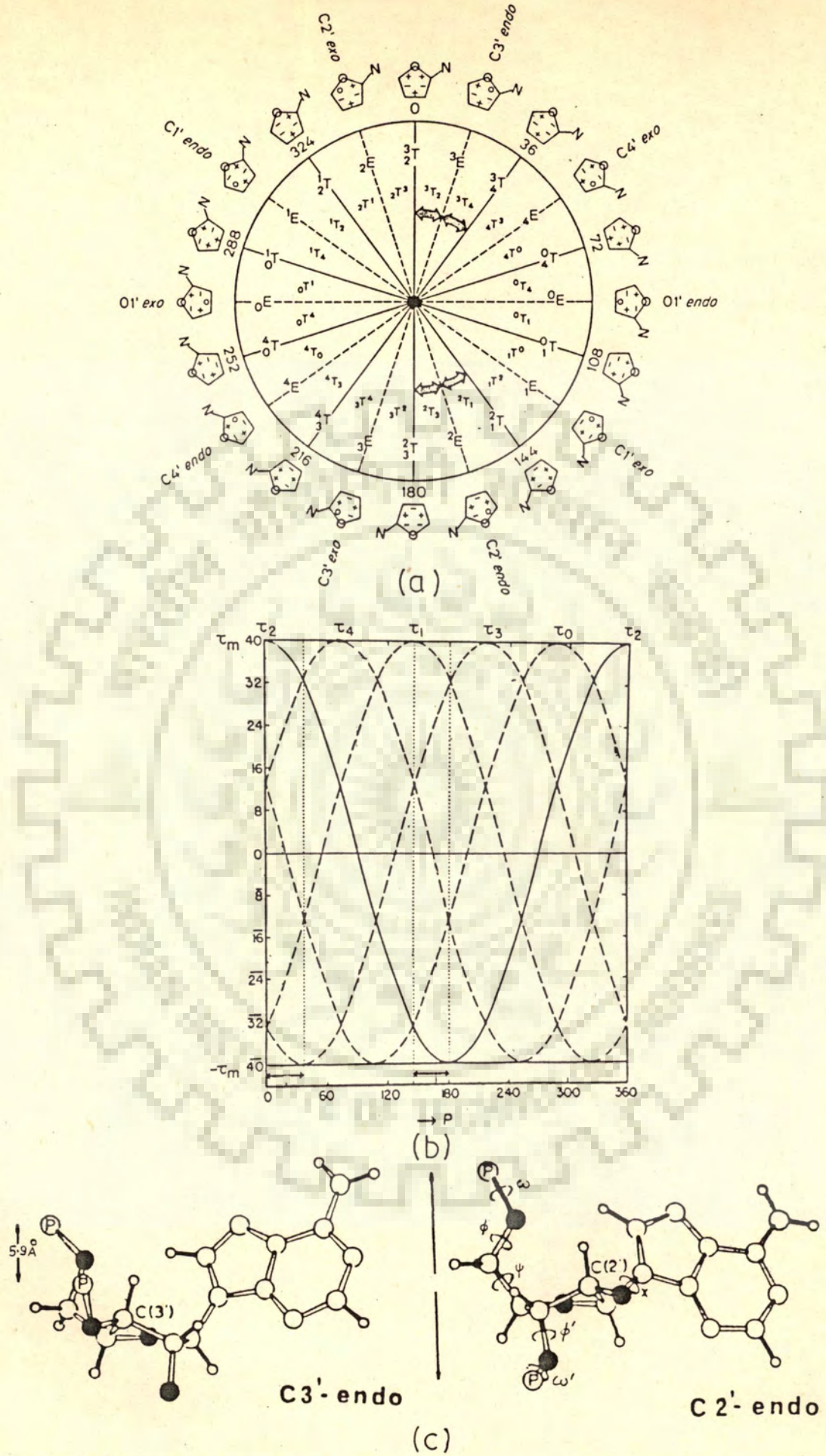


Fig. 1.2 **(a)** Classical Conformation of the furanose ring as a function of P **(b)** Values of five torsion angles τ_0 to τ_4 in the furanose ring as P goes from 0° to 360° **(c)** Preferred conformations C3'-endo and C2'-endo of sugar pucker.

description. In C2'—endo geometry the C2'—atom is on the same side of the C4' - O1' - C1' plane as the C5' atom. Similarly C3'—endo pucker refers to a geometry in which the C3' atom is on the same side of the plane as the C5' atom. The barrier height for the inter conversion of these two forms is estimated to be 4.0 kcal/mole [51].

The rotational position of the base relative to the sugar is sterically restricted and two conformational states, namely anti and syn, are preferred (see Fig.1.1). For χ_{CN} as anti there is no particular steric hindrance between sugar and base but in Syn, the bulky part of the base is located over the sugar giving rise to close inter atom contacts.

Variety of physico-chemical techniques and theoretical calculations have shown that the geometries about various bonds are confined to some well defined conformational domains. Major degrees of conformational flexibililty in nucleic acids arises from the internal rotations about phosphodiester bonds. An analysis of available data show that the most preferred conformations of nucleic acid backbone correspond to a set of dihedral angles $\Phi' = 220°$, $\omega' = 290°$, $\omega = 290°$, $\Phi = 180°$, $\psi = 60°$, $\psi' = 83°$. The resulting structure with these conformations is a regular right-handed helix.

Various studies on DNA/oligonucleotides in solid state and in solution have shown the existence of primarily three families of nucleic acid, double -helices These are the right-handed

B—family with antiglycosidic bonds and C3'—exo/C2'—endo sugar pucker; the right—handed A—family with antiglycosidic bonds and C3'—endo sugar pucker; and the left—handed Z—family with alternating syn-anti glycosidic bonds and alternating C3'—endo-C2'—endo sugar pucker. The characteristic features of these are summarised in Table 1.2. The RNA - RNA duplexes always exist as A—type and right—handed because the 2'—OH group prevents a B—type conformation. DNA-RNA hybrids also adopt an A—type structure. Native DNA is found to exist in B-DNA form.

The left—handed structures have been observed only in alternating GC oligodeoxynucleotides. Cooperative transitions between families can be induced by salt, alcohol or humidity changes.

The fibre diffraction, single crystal diffraction and solution studies on nucleic acids/oligonucleotides have confirmed the existence of variety of conformations discussed above. These have been reviewed in literature [62, 123] extensively. Recent studies by Frechet et al. [40], Feigon et al. [38], Hare et al. [56], Pardi et al. [104], Broido et al. [15] on specific deoxyoligonucleotide in solution by two dimensional NMR have yielded unambiguous analysis of their conformation. The structural analysis of some oligonucleotides which are important from the view point of recognition of nucleic acid base sequence by certain properties in a highly

Table 1.2 : Characteristics of the three families A,B,Z of Double Helical DNA (62)

	A-DNA		B-DNA		Z-DNA	
Helix sense	right-handed		right-handed		Left-handed	
Glycosidic bond	anti		anti		syn(deoxyguanosine) anti(deoxycytidine)	
Sugar pucker	C3'-endo		C3'-exo/C2'-endo		C3'-endo (deoxyguanosine) C2'-endo (deoxycytidine)	
	A	A'	B	C	Z	Z'
Bases per turn	11	12	10.4-10	$9\frac{1}{3}$	12 (6 dimers)	12
Rotation per base	32.7°	30°	34.6°-36°	38.6°	-60°	-60°
Helix diameter (nm)	2.3	2.3	1.93	1.92	1.81	1.84
Rise per residue (nm)	0.2	0.3	0.34	0.33	0.74/2	0.76/2
Helix pitch (nm)	3.09	3.6	3.38	3.1	4.46	4.57
Base pair tilt	16-19°	10°	-6°	-8°	-7°	-9°

specific manner have also been solved recently by 2D NMR. The conformational study at Bam Hl cleavage site [79] has been carried out for a self complementary dodecanucleotide d-GGATCCGGATCC by Govil and coworkers. It has been shown that this dodecanucleotide assumes a predominantly B-conformation with sequence dependent changes along the chain. The recognition site of Bam Hl shows a distinctly different geometrical environment. The sugar rings of Gl and G7 assumes a C3'—endo geometry while the rest of sugars possess C2'-endo geometry. All nucleotides are found to be in anti conformation.

Similarly to study the cleavage site of another restriction enzyme ECORI the structure of dodecanucleotide d-GAATTCGAATTC has been determined [23]. Most of the sugar rings adopt an unusual O1'—endo geometry, all glycosidic dihedral angles are in anti domain. The AATT segments A2-T5 and A8-T11 show better stacking compared to the rest of the molecule. There are important deviation in the central TCG portion which is known to show preference for DNase I activity and between G1-A2 and G7-A8 which are cleavage points in ECORI recognition sequence. The sugar puckers for G1 and G7 are significantly different from the rest of the molecule. Further in the three segments mentioned above the sugar phosphate geometry is such that the distance between protons on adjacent nucleotides are much larger than those expected for a right—handed DNA. Such crevices in the DNA structure

may act as hot points in initiation of protein recognition.

ELEMENTS OF PROTEIN STRUCTURE

All proteins and enzymes are linear polymeric molecules whose monomeric units are α-amino acid residues. The amino acids except proline have the characteristic chemical structure NH-CHR-COOH. The carbon atom which is bonded to -NH and -COOH groups is designated the α-carbon and the carbonyl carbon atom is labelled as C'-carbon. The nature of the side chain R distinguishes one amino acid from another. Based on the polarity of the side chains the amino acids are classified as hydrophobic, polar, positively charged and negatively charged. All amino acid except glycine have an asymmetric α-carbon. Of the two stereo isomers only L-amino acids are found in naturally occurring proteins.

Condensation of the amino group ($-NH_2$) of an amino acid with the carboxy group -COOH of another amino acid results in the formation of the peptide O=C-N-H bond. Delocalisation of the π-electrons give partial double bond character to the C'-N bond. Hence six atoms C^{α}, N, H, C', O and C^{α} lie in the same plane and the rotation about C' - N bond becomes restricted. The protein backbone can be viewed as a series of amide plane joined by tetrahedral carbon atoms. In spite of the rigid planar peptide groups, rotations about the C-N and C-C bonds are possible. The conformations about these bonds are defined by two dihedral angles, Φ , about

C-N bond and, ψ, about C-C bond. A list of all the Φ_i, ψ_i for all the residues defines the conformation of the polypeptide backbone. All the residues of a fibrous protein will have the same Φ, ψ values. The residues of a globular protein, on the other hand, have a wide range of Φ, ψ values. The actual values of Φ, ψ along the backbone depend on the primary structure of protein. [See for example ref 128].

The nonbonded interactions between the various atoms limit the range of permissible conformations with respect to Φ and ψ. Considering that two atoms may not approach closer than a certain limiting distance Ramachandran et al. [119] constructed a map showing the allowed conformation of Φ, ψ. It is satisfying to see that the structures corresponding to α-helix, β-sheet, collagen etc. fall within the allowed regions of Ramachandran's map. The α-helix characterised by 3.6 residues per turn, gains its additional stability from the intra—chain hydrogen bonds between the N - H and the C = O groups of successive coils. In this structure, $\Phi = -57°$ and $\psi = -47°$. In the β—sheet structure, the individual chains are rather extended and the neighbouring chains running parallel or antiparallel are hydrogen bonded. In collagen three left—handed helices are held together by inter—chain hydrogen bonds. In contrast to the regular uniform structure of the fibrous proteins, the globular proteins have diverse 3-dimensional structure. These structures do contain stretches of α—helices, β—sheets

and fold to result in a well defined tertiary structure dependent upon its primary structure.

PROTEIN-NUCLEIC ACID COMPLEXES

DNA BINDING PROTEINS

Since DNA acts as repository of genetic information most of its interactions with the proteins facilitate the storage, replication or repair of DNA. A representative list of such proteins involved in (a) repair and replication is - ligases and repair proteins, nucleases and excision enzymes, gyrases, DNA polymerases (b) transcription - RNA polymerase, cyclic AMP receptor protein, repressors, rho factors which terminate transcription (c) nucleases - restriction endonucleases, exo- and endonucleases (d) recombination rec A protein and (e) package - histones, protamines, virus condensation proteins, etc.

A broad range of specificity is observed in protein nucleic acid associations. Those involved in package and repair are less specific to the base sequence. The coulombic interactions between the phosphate groups and the basic amino acid residues of the proteins are assumed to be the major source of binding in these cases. In contrast proteins which facilitate and regulate the transcription are very specific to the nucleotide sequence. For example, lac and lamba repressor proteins bind with high degree of specificity to the operator regions which are but few residues on a

long genome of about million base pairs. The lac repressor is cleaved between residues 51 and 59 depending upon the proteases used [44], the core protein remains a tetramer with 300 amino acid residues per protomer from which the head piece (51 to 59 residues) can be separated. The head piece still binds to lac operator sequence in a specific way [101]. The λ repressor is also cleaved [103] by proteases and yields fragments of nearly equal size and the N-terminal fragment (92 residues) of repressor is shown to consist of five α-helices and an extended arm. The helix 3 protrudes out of the molecular surface. It is suggested [103, 125] that helix 3 interacts with the major groove of DNA while helix 2 which is close to helix 3 interacts with nucleic acid backbone. Structural information in solution has been obtained from neutron [22] and X-Ray scattering [112].

Many restriction endonucleases and modification enzymes are capable of recognising the base sequences. These are tabulated by Smith [132] and by Roberts [122]. In such cases the bases are found to be partially or fully palindromic [20]. The exonucleases and some endonucleases, gene 32 protein of bacteriophage fd, DNA melting proteins, etc. recognise the strandedness of the helix. For such a recognition the binding constant must be 5 - 6 orders of magnitude greater than the nonspecific binding. Crystal structure of gene 5 protein from bacteriophage fd, a single strand binding protein and its complexes with oligodeoxynucleotides

has been solved [91, 92]. The gene 5 protein has a three stranded antiparallel β-sheet and two distinct antiparallel β- loops joined by short segments of extended polypeptide chains. It contains a long groove beneath the three-stranded sheet which could be the DNA binding site. The interior of the groove contains basic amino acid residues while the edges bear a number of aromatic side chains which could be involved in stacking interactions with the bases as suggested by NMR experiments in solution [26].

The crystal structure of Staphylococcal nuclease has been solved in the absence of and in the presence of the inhibitor thymidine-3', 5'-biphosphate (pTp) [30]. It is found that thymine ring of the inhibitor pTp fits into a hydrophobic pocket. The tyrosyl ring of Tyr 113 is involved in weak stacking interactions with the thymine ring. The guanidinium ions of Arg 35 and Arg 87 each form a pair of hydrogen bonds with the 5'-phosphate. Such pairs of hydrogen bonds between guanidinium and phosphate ions have been observed in model systems [29]. The two side chains of Arg 35 and Arg 87 not only bind to the phosphate but also form additional hydrogen bonds with other groups in the enzyme thus providing a very precise location of the 5'-phosphate with respect to the enzyme.

The cro protein which binds to the same DNA sequences as the λ CI repressor consists of 66 amino acids as seen

by crystal diffraction to a resolution of 0.28 nm [4]. The structure consists of 3 strands of antiparallel β-sheet (residues 2-6, 39-45 and 48-55) and three α-helices (residues 7-14, 15-23, and 27-36). The four carboxy-terminal residues are disordered in crystal and there are four monomers per asymmetric unit. An extended C-terminal tail interacts with the corresponding tail of another monomer and dimers interact to form a tetramer. The α-helix from Gln 27 to Ala 36 appears to interact with large groove of DNA. The corresponding helix in the other subunit of the dimer is parallel and centre-to-centre distance of two helices is 3.4 nm. The two helices have the correct respective position to be accommodated within the successive major grooves of the DNA on the same face of the B-right-handed double helix. Close to the two fold symmetry axis of the dimer a pair of antiparallel β-sheet strands (Glu 54-Val 55-Lys 56 of each monomer) could fit the minor groove between the two major grooves occupied by the α-helices. The C-terminal residues, including Lys 62 and Lys 63, could also interact with DNA. The structure explains known properties of cro - λ operator interaction, protection against methylation, size of binding site etc.

The catabolite gene activator protein, (CAP), complexed with cyclic AMP has been crystallised and its structure is determined [91]. Cyclic AMP binds to the larger amino-terminal domain (~135 amino acids) while the smaller carboxy-terminal

domain (~65 amino acids) is presumed to bind DNA. The main feature of the model proposed to explain how CAP dimer binds to DNA is that two 2.2 nm long α—helices (one in each subunit) fit into successive major grooves on one face of a left—handed B—DNA double helix. The structural complementarity between the two molecules extends over 18-20 base pairs (6.5 - 7 nm) of DNA which is the size of the region protected by CAP dimer in solution. Neither a right—handed B—DNA or Z—DNA could accommodate the two α-helices which are nearly parallel to each other and inclined at ~65 - 70° with respect to line connecting their centres and whose helix axes are separated by 3.4 nm. The model suggests that CAP binding to DNA might induce a right—to left— helix transition and destabilise neighbouring regions involved in the control of transcription initiation.

A comparision of amino acid sequence of various regulatory proteins suggest that the sequence are quite homologous [126]. Currently, it is believed that the involvement of two α-helical fold, similar to $\alpha_2 - \alpha_3$ in cro protein, in binding to the major groove of DNA is a common feature of nucleic acid-protein recognition. The sequence specificity in these models arise from the formation of hydrogen bonds between the amino acid side chains and the bases/base pairs. However, these models are not confirmed experimentally. Klug et al. [72] suggest that the backbone conformation of DNA is dependent on the base sequence and this variation

in the backbone geometry is recognised by the proteins. The alternating phosphodiester conformation in d-ApTpApT and the enhanced affinity of lac repressor to poly (dA - dT) than that for calf thymus DNA support this hypothesis. The mechanism involved in finding the operator site by proteins is an intriguing question in the specific recognition. However, Berg et al. [11] have shown that the lac repressor first binds to non-operator site and then slides along the DNA by a one-dimensional random walk.

RNA BINDING PROTEINS

The involvement of RNA in a great variety of functions is reflected in its association with proteins as well. A representative list of RNA binding proteins is (a) those involved in polymerising and repair enzymes are RNA polymerase, RNA ligase, reverse transcriptase, replicase, (b) proteins interacting with tRNA - proteins cleaving rRNA precursors, rRNA modifying enzymes, structural proteins of ribosomes, initiation and elongation factors, peptidyl transferase (c) proteins binding to mRNA - maturation enzymes, poly-adenylation enzymes, hydrolytic enzymes (d) Proteins interacting with tRNA-maturation and modification enzymes, initiation and elongation factors, aminoacyl- tRNA synthetase, G factor for translocation and (e) nucleases ribonuclease, exo- and endonucleases, hydrolases, etc.

Bovine pancreatic ribonuclease (RNase A) and RNase S obtained from proteolytic cleavage of RNase A between residues 20 and 21 have been crystallised [121, 146] in the absence and the presence of substrate inhibitors such as mononucleotides of modified dinucleoside monophosphate. The specificity of RNase A with respect to pyrimidines appears to reside in the formation of hydrogen bonds between the protein backbone or side chains with the pyrimidine bases. The capability of the side chain of Threonine 45 to behave either as a proton acceptor or as proton donor in hydrogen bonds allows binding of either cytosine (amino group as donor) or uracil (carbonyl group as acceptor). This degeneracy in recognition involves only a rotation of the -OH group of threonine around the C-O bond. In the complexes formed by RNase S and UMP, CMP, UpcA another hydrogen bond links the O(2) of pyrimidine base (C or U) to the backbone nitrogen of Thr 45. In crystal structure of RNase S-$C_{2'p5'}A$ complex [146] the binding of adenine involves hydrogen bonding to several groups of the enzyme (Gln 69, Asn 71, Glu 111).

Several features of RNase S-inhibitor complexes are shared with complexes of dehydrogenases with coenzyme NAD^+. The phosphate groups do not seem to be engaged in strong charge - charge interactions. The adenine moiety in NAD^+ appears to interact with the dehydrogenase via nonpolar groups whereas adenine in $C_{2'p5'}A$ forms several hydrogen bonds with RNase S. Although NAD is not a part of nucleic acid, it may be noted

that glyceraldehyde -3-phosphate dehydrogenase as well as lactate dehydrogenase have been shown to bind to nucleic acids and to exhibit a selectivity towards single—stranded structures. A study of oligonucleotide binding has shown that these dehydrogenases have a strong preference for dinucleotides containing adenine [112].

In contrast to eukaryotic ribonuclease A the prokaryotic ribonuclease T1 extracted from the fungus Aspergillus oryzae exhibits [111] very specific recognition for only one base guanine. The crystalline complex between this enzyme and the inhibitor 2'—guanylic acid (2' - GMP) provides for a detailed picture of combined hydrogen bonding and stacking interactions involving protein main- and side—chain atoms which envelope and bind the guanine base [111]. It is found that side chain of Tyr 45 swings over and stacks at a 3.5Å distance with G, an interaction certainly favouring purine over pyrimidine bases and guanine is in syn orientation not easily accessible for pyrimidine bases.

Tobacco mosaic virus, (TMV) has been the subject of structural investigations for long. The TMV protein unit contains 158 amino acids and 2140 subunits interact with a single—stranded RNA 6400 nucleotides long [13]. Three nucleotides interact with each protein subunit. The proposed structure [63] involves interactions of two invariant residues Asp 115 and Asp 116 with the 2'—hydroxyl group of the two

outer riboses in each nucleotide triplet. The central base may form hydrogen bond with Ser 123. The formation of salt bridges between phosphates of one base triplet and the arginine residues 90 and 92 of the neighbouring protein subunit has also been suggested. The region of viral RNA which inserts into disc to initiate assembly of the virus has in its centre a sequence with every third residue as guanine. This feature might be important for the selectivity of protein-RNA interactions which initiate virus assembly. The crystal structures of some of the tRNA—synthetases and their complexes with cognate tRNA are under investigations [46].

NATURE OF INTERACTION BETWEEN FUNCTIONAL GROUPS IN PROTEIN-NUCLEIC ACID ASSOCIATIONS

Due to the polymeric nature of nucleic acid, and proteins, multiple sites should be involved in their binding to make it strong and effective. As in any enzyme substrate interaction structural complementarity in the protein nucleic acid association is a pre-requisite. Inspite of the complex structure of these macromolecules involved, the intermolecular forces involved are different from those engaged in any type of molecular complex. The interaction energy between two molecules may be evaluated by supermolecule approach in which the total energy of the supersystem is calculated and the energies of the two subsystems are subtracted. Alternately, interaction may be treated as a perturbation

of the isolated subsystems and the total interaction energy can be split into several terms : coulombic interaction between the charge distribution of the two molecules. polarisation, dispersion, exchange repulsion and charge transfer [25]. In calculations performed for interaction between nucleic acid bases, it is found that the main contribution to the stability of stacked complex arises from the dispersion forces whereas hydrogen bonding is characterised by a preponderance of the electrostatic term [82].

The interactions between functional groups have been characterised in terms of four major types - electrostatic, hydrogen bonding, stacking and hydrophobic interactions. Although this classification is a little bit arbitrary it has the merit of distinguishing the contribution of different types of amino acid side chains to the stability of the complexes. The borderlines between various interactions are not very sharp, for example coulombic interactions involve hydrogen bonding as well. Similarly stacking interactions are basically hydrophobic. Much of our knowledge is acquired from the studies on model systems. In the following sections these are discussed very briefly except for stacking interactions which are discussed at length.

1. Electrostatic Interactions

Most of the protein—nucleic acid complexes involve electrostatic bonds between positively charged lysyl, arginyl

or histidyl side chains with negatively charged phosphates. The binding of DNA with histones has been studied extensively [39] and similar forces are found to be operative in binding of protamines which are rich in Arg, with DNA in higher organisms. There are several other instances where they play important role [2, 116]. Inspite of their non directional nature coulombic interactions can contribute to the specific recognition of nucleic acids by proteins. For example Histone-H1 rich in Lys^+ is specific for A-T base pairs [14], Poly (Lys) is specific for A-T rich regions of DNA while Poly (Arg) recognises G-C rich regions [83]. The binding of basic proteins to nucleic acids can alter the secondary structure of the later. Histones stabilise the regular helical structure [14] whereas pancreatic ribonuclease destabilises the DNA [71].

Some of the model systems have also been investigated [59] to provide evidence for the possible role of metal cations by forming a bridge by chelation between negatively charged amino acids and phosphate backbone. Cu^{2+} and Zn^{2+} have been used for the binding of Poly (Glu) and Poly (Tyr) with polynucleotides. Crystal structure of staphylococcal nuclease-inhibitor complex provides an example of metal cation (Ca^{2+}) acting as a bridge between the enzyme and the inhibitor [49]. Gresh and Pullman [53] have studied the interaction of ammonium and guanidinium ions with phosphate anion by SCF abinitio and empirical procedures. Vasant

and Govil [76] have made interaction energy calculations for charged amino acids with a view to understand binding affinities and relative stabilities of their complexes with nucleic acids. They have shown that Lys^+ and His^+ form more stable complexes with RNA than with single-stranded DNA whereas Arg^+ has higher affinity for DNA than for RNA. The affinity decreases in the order $Lys^+ > His^+ > Arg^+$. Determination of the number of electrostatic bonds in complexes and more examples of specific recognition by coulombic interactions is discussed elsewhere in literature [61].

2. Hydrogen Bonding Interactions

Hydrogen bonding may occur between amino acid side chains of Asp, Asn, Glu, Gln etc. or the peptide backbone with nucleic acid constituents particularly the base/base pairs. For proteins the groups OH (Ser, Thr, Tyr), SH (Cys), -S-(Met), COOH or COO^- (Glu, Asp), $CONH_2$ (Gln, Asn, peptidic backbone), NH_2 or NH_3^+ (Lys, Arg), NH (Trp, His), N (His) and for nucleic acids the groups $\geqslant$ N(A, G, U), > N-H. (U, G), $-NH_2$ (A, C, G), C = O (U, C, G), -OH(ribose), PO_4^- (Phosphate backbone) may participate in hydrogen bonding. Since the polarities of these groups vary widely, very different association constants should be obtained for the large number of possible interactions which could occur. (Fig. 1.3)

All polar amino acid side chains are able to form atleast one hydrogen bond with each of the four bases since

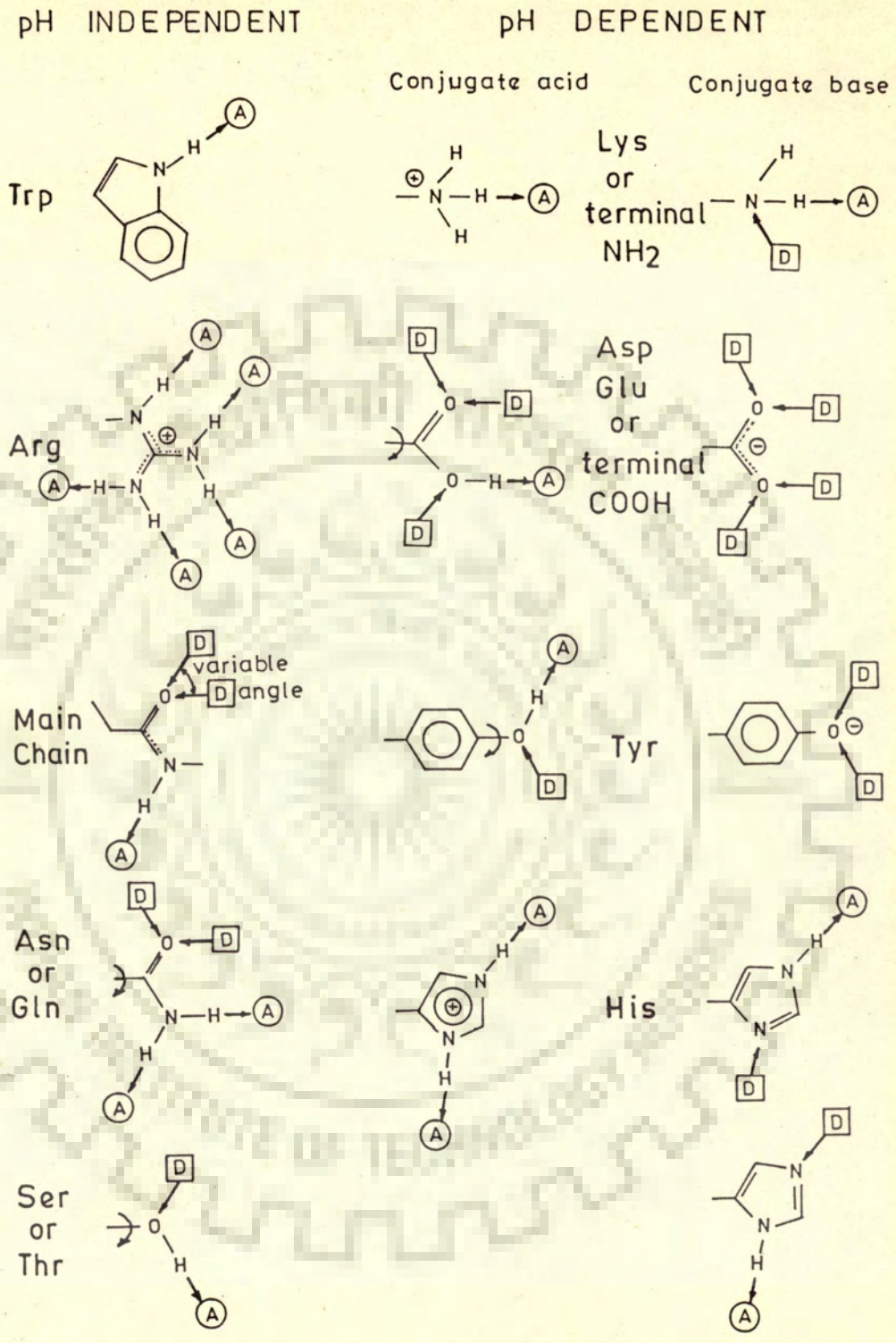


Fig. 1.3 Hydrogen bonding possibilities of protein side chains and backbone with acceptor (A) or donor (D) groups.

each one of them possesses donor and acceptor groups of hydrogen bonds. This peculiarity has led some scientists believe [16, 60, 61] that a single hydrogen bond is in adequate for uniquely identifying any particular base or base pair, however, with two hydrogen bonds base pair recognition may be achieved. Unionised side chains containing acid, amide, alcohol or sulphhydryl functions can form two hydrogen bonds with paired or unpaired purines and only with unpaired pyrimidines. Even though neutral Asp and Glu, Asn and Gln side chains can form two hydrogen bonds with purines in A-T and G-C base pairs a discrimination arises from the groove where interaction can take place. A-T will be recognised from the major groove and G-C from the minor groove of DNA. The highest specificity of interaction can be obtained with ionised side chains. The side chain of Arg has been proposed [60, 61] to form two hydrogen bonds with unpaired guanine or cytosine or with paired guanine. A model has been proposed where the positive charge of Arg side chain is neutralised inside the protein structure by a carboxylate anion whose two oxygen atoms could form hydrogen bonds with two NH groups of Arg. A sequence Arg-Glu would be more favourable than Arg-Asp for steric reasons [61] and interaction would be favoured in the Arg-Glu-Lys sequence where the ionised side chain of Lysine can interact with a phosphate group [61]. Sequences such as Arg-Glu, Arg-Glu-Lys, etc. are found in proteins such as lac repressor, λ repressor and some ribosomal proteins. Ribose or phosphate

groups can also form hydrogen bonds with the polypeptide backbone or side chains. Several other models of interaction between bases and amino acid side chains or the peptide backbone have been proposed [16, 64].

Crystallographic studies have shown existence of hydrogen bond in several nucleic acid complexes with proteins. Jain and Sobell [67] have shown the existence of strong hydrogen bond between amino group of guanine and carbonyl oxygen of threonine residue while a weaker hydrogen bond connected the nitrogen atom N3 of guanine with NH of threonine in actinomycin-deoxyguanosine complex. In crystals of thymine linked via a propyl bridge to the indole side chain of tryptophan Voet and Bunick [143] showed that the O2 oxygen atom of thymine is hydrogen bonded to the NH of indole. Hydrogen bonds between N1 of neutral adenine and peptidic NH [105] or N1H of protonated adenine and peptidic C = O [106] were found in derivatives of adenine substituted by a peptidic bond on the N6 of amino group. Hydrogen bonds have also been found between phosphate anion, on one hand, and hydroxyl group of Tyr 33 and amino group of Arg 52, on the other hand, in the complex of phosphorylcholine with mouse immunoglobulin Fab. [129].

Evidence of hydrogen bonding can be obtained from UV, IR, absorption, Raman or NMR spectroscopy. The A-H bending frequency increases, the A-H streching frequency

decreases, the bandwidth and band intensity of the A - H stretching vibration increases in the IR absorption spectra and Raman emission spectra on hydrogen bonding. In NMR spectra, it leads to a downfield shift of proton resonances which are directly involved in hydrogen bond formation. Difference absorption measurements [81] have shown values of association constants that are as large as 22000 M^{-1} for hydrogen bond between Asp, Glu and 9-ethyladenine in cyclohexane solvent. Several results on hydrogen bonding between bases and neutral/ionised side chains of amino acids in water and nonpolar solvents exist in literature [62]. Peptide backbone also participate in hydrogen bond formation as shown for binding of uracil to peptide backbone in chloroform forming two hydrogen bonds: N3-H---O = C and C4 = O---- H-N [98). The association constant is of the same order of magnitude as the association constant for A-U base pair formation and this result suggests a complementarity between uracil and peptidic backbone. The hydrogen bond energies are in order of magnitude smaller than the coulombic interaction energies between basic amino acids (Lys^+, Arg^+, His^+) and phosphate groups of nucleic acids. Energy calculations have shown [77] that the stabilities of the complexes of Asn, Gln, Asp, Glu with bases are in the order G - X > C - X > A - X > U - X or T - X and G - C—X > A - T (U) - X where X is one of these amino acid residues. Further Glu^- and Asp^- can recognise guanine in single-stranded nucleic acids, Arg^+ can recognise G - C. base pair from A - T base pair in double-stranded structures.

3. Hydrophobic Interactions

These interactions are expected between the aliphatic amino acid side chains of Val, Leu, Ile, etc. and the bases of nucleic acids. Binding of proteins to nucleic acids is accompanied by large positive entropy changes and possibly by large heat capacity changes. Among the possible sources of entropy and heat capacity changes in processes involving proteins, conformational, hydrophobic and vibrational effects seem likely to be of greatest importance [137]. If hydrophobic interactions, resulting in the release of bound structured water were important, one would expect the association to be favoured at high salt concentration since an increase in electrolyte activity would reduce the water activity and the release of bound water into the solution of lower water activity would drive the hydrophobic association towards completion. This is not experimentally observed and it is, therefore, difficult to estimate the contribution of hydrophobic interactions to the process of protein-nucleic acid associations when ionic interactions dominate the process. However, in some cases, such as the binding of ribosomal protein S1 to oligoribonucleotides, there is a small increase of the binding constant when NaCl concentration increases [35] which might indicate that the association reaction is driven by water release. A highly structured water network has been recently described in crystals of

d-GpC—proflavin complexes [97] though the role played by these structured water molecules in protein binding is difficult to assess in solution studies. The high entropy of activation for the synthesis of Ile - tRNA has been attributed to the loss of structured water [84]. This might result from a conformational change which modifies the accessibility of water—organising residues at the interface between water and the enzyme—substrate complex. Loss of structured water could occur at some distance from the active site in a coupled reaction.

Changes in specific protein—nucleic acid association induced by organic solvents [89, 45] are difficult to infer because they may act on the stability of the nucleic acid and facilitate a local change in conformation. They might also change the conformation of the binding site of the protein and therefore, alter the specific/non—specific binding ratio as well as the kinetics of catalytic step if the protein possesses an enzymatic activity. There is evidence for the involvement of hydrophobic contacts in the lac repressor-lac operator interaction. Using synthetic lac operator fragment, an important hydrophobic contact between the lac repressor and the 5-methyl group of thymine has been found to exist. The presence of a hydrophobic region of the repressor might also explain the enhanced methylation rate of this A - T and neighbouring G - C base pairs due to an increased local concentration of dimethyl

sulphate [48].

4. Stacking Interactions

The stacking interactions, similar to those found in nucleic acids could occur between aromatic side chains of Trp, Tyr, Phe and His and the bases or base pairs of the nucleic acids. Many aromatic molecules or intercalated complexes are characterised by unwinding of the nucleic acid double-helix and insertion of aromatic ring in the space between two base pairs. The structure is characterised by mixed C3'-endo (3' - 5') C2'-endo sugar puckering, though this is not always the case [130] and may be accompanied by a local bending or kinking of the double helix [136].

Aromatic amino acids do not have the same size as the aromatic dyes which have been shown to intercalate in double helices and which have a size comparable with that of a base pair. The indole ring of tryptophan has the same size as purine. The phenol and phenyl rings of tyrosine and phenylalanine are comparable to a pyrimidine. Therefore, it is not likely that they intercalate between base pairs, on the contrary, they might partially stack with bases on one strand, leading to a bending of the double helix or with base pairs, eg. in the minor groove, inducing a kink in the duplex structure. The situation might also be viewed as trapping of pre-existing bends or kinks of

the double—helix by the aromatic residue through partial stacking. Irhediamine which does not possess an aromatic planar ring system could partially stack with bases [134]. A partial insertion model has also been proposed for the binding of a steroidal diamine (dipyrandium) with Poly (dA-dT) on the basis of NMR studies [110].

Evidence for stacking interactions between aromatic amino acids and nucleic acid bases come from nuclear magnetic resonance, fluorescence and circular dichroism studies of oligopeptide binding to polynucleotides and nucleic acids. Gabbay et al. [42] had shown that aromatic amino acids of peptides and peptide amides partially intercalate between the base pairs of salmon testis DNA. These studies have been reviewed in literature [62].

A. NMR Studies

Earlier experiments by proton NMR have shown that aromatic residues of oligopeptides form stacked complexes with nucleic acid bases [42, 58, 33, 89]. Stacking was found to be favoured in single-stranded nucleic acids, polynucleotides (such as Poly A) or UV irradiated/denatured DNA rather than double helical native DNA [33, 89]. The studies on single strand binding proteins gene 5 protein of phage fd [26,27] and of phage M 13 [43a] by NMR have shown that tyrosyl residues Tyr - 26, 34 and phenylalanine residue Phe 13 and His 64 are involved in binding to single stranded oligonucleotides

e.g. $(dA)_8$ through stacking interactions with bases.

Photochemically induced nuclear dynamic polarisation (CINDP) studies have shown that two out of four tyrosine residue in the N-terminal head piece of the lac repressor are protected in the complex with oligo d(AT) [17]. The measurements on isolated headpiece show that Tyr 7, 17, 47 and His 29 are accessible for interaction but Tyr 12 is much less accessible. But in the complex with oligo d-(AT), Tyr 7, 17 and His 29 are no longer accessible to the photo-sensitiser. These are therefore involved in contact region. The involvement of Tyr 7, 17 in the interaction of N-terminal region of the lac repressor with DNA has also been inferred from chemical modification experiments [37, 1]. Three out of the five Tyr residues of gene 5 protein from phase M 13 also give rise to photo-CINDP effect. These are the three residues accessible on the surface of protein. In complexes with the tetranucleotide d-CGCG these tyrosines are no longer accessible to the excited flavin and their photo-CINDP signal disappears [43b].

B. Fluorescence Studies

Aromatic amino acids are fluorescent at room temperature with a quantum yield that depends on the immediate environment in the protein structure [85] whereas that for nucleic acid bases is two orders of magnitude lower than say that of tryptophan. The quenching of tryptophan fluorescence when

the indole ring is stacked with nucleic acids bases [94] has been ascribed to an electron transfer in the excited state from indole to purine or pyrimidine bases. It has been shown that this could be due to a conformational change or the protein which brings a protein quenching group in the vicinity of a tryptophyl residue, or more generally, for proteins which bind to single-stranded nucleic acids, point to the involvement of a more general mechanism which could be the partial stacking of one or more tryptophyl residue with nucleic acid bases. Since the fluorescence decay in these nanosecond time range was not markedly affected when compared to the free peptide, it was concluded that the Trp is involved in stacking interaction with complete quenching of the stacked indole rings. A sequence dependence for tryptophan stacking with bases has been observed [88] in single-stranded polynucleotide wherein the stacking between adjacent bases decrease in the order UU > UA, AA > AC, CC.

Helene et al. [57] have shown that fluorescence of tyrosine is quenched in stacked complexes with bases which is probably due to electron transfer in the case of pyrimidine bases but might involve energy transfer in the case of purines. In complexes of staphylococcal nuclease with the inhibitor pTp the fluorescence of the single Trp residue is not affected whereas, that of Tyr residue is quenched [32]. The crystal structure of the complex reveals that Tyr 115 is very close and stacked with thymine. The Tyr fluorescence of the gene

5 protein from phage fd is quenched upon binding to single stranded nucleic acids [118]. This quenching, however, is not a proof of stacking. The fluorescence of Phe is difficult to detect in a protein and is observed only in special cases wherein Tyr and Trp residues are not present. It has been found that fluorescence of peptides Lys-Phe-Lys and Lys-Phe (NH_2) is quenched upon binding to single—stranded Poly-nucleotides and double stranded DNA [89]. Helene et al [140, 34, 21] have studied the role of tryptophyl residues in the binding of Gene 32 Protein from phage T4 to single-stranded DNA and found that three tryptophan out of total five tryptophan are not able to bind to double stranded DNA after irradiation. They utilized luminescence properties of aromatic molecules to detect tryptophan in vicinity of nucleic acid.

C. Theoretical Studies

Inspite of the large number of studies on stacking interactions involving aromatic amino acids by NMR, CD, fluorescence, viscometry, etc. very little is known about the intrinsic stabilities of the stacked complexes of aromatic amino acid residues with the bases or base pairs. Details about the relative orientation of the two components, that is the overlap geometry of aromatic amino acid within two bases/base pairs are not known. Stacking energies of Tyr, Trp, His, Phe with bases and base pairs have been calculated [78] by second order perturbation theory. The major contribution to the

stacking energy is found to arise from dispersion energy terms. It has been shown that among pyrimidines cytosine forms a more stable complex whereas in purines the guanine forms a complex which is more stable than that with adenine. There is a considerable overlap of the aromatic rings of the base and amino acid in the stable configuration.

Scope of Thesis

So far we have discussed the structure of nucleic acids and protein/protein fragments, general features of protein—nucleic acid interactions and a literature survey of these interactions (a representative one and not an exhaustive one). We find that today our knowledge of interactions between proteins and nucleic acids is still rather limited. The main problem is that these systems are all rather complex and require simultaneous observation of two associated macro molecules. Spectroscopic methods are except for a few favourable cases inadequate and crystallisation of protein—nucleic acid aggregates is difficult. In order to circumvent the difficulties inherent in these systems, model compounds have been investigated using both theoretical and experimental approaches. One major concern has been learning about the specificity of recognition between the four nucleic acid bases and the 20 amino acid chains, using both monomeric constituents of both partners as well as polymers, in some cases polymers of synthetically available peptides or nucleic

acids. The obvious advantage of this approach is the possibility to investigate interactions of DNA with peptide side chain with controlled amino acid composition, sequence, chain lengths and predictable conformation which results in a considerable simplification of the systems as compared to the natural protein - DNA systems. Further, such structures provide insight into the requirements of protein–nucleic acid interactions. This kind of approach has particularly become more relevant with the potentialities of some of the recent experimental techniques, namely two dimensional nuclear magnetic resonance which probe directly into detailed conformational features in solution. With this in mind, we have undertaken a study of stacking interaction involving aromatic amino acids Tyr, Trp and Phe with deoxyoligonucleotides containing C and G bases.

The oligonucleotides containing bases C and G were preferred over A and T bases only with an idea that a left-handed DNA structure may be seen in such complexes. Specific oligonucleotide sequences, that is, d-CpG, d-GpC, d-CpCpGpG and d-GpCpGpC were selected to investigate if there was a preference for dG-dC or dC-dG site and also if an alternating base was preferable over other sequences. Most of the work was carried out by 2D-NMR methods in solution. Chapter–2 details out the materials used and the methods used in analysis of data.

Chapter—3 to—6 deal with results and discussions of NMR data on interacting molecules. Chapter—7 deals with theoretical energy calculations on intercalation of Trp, Tyr and Phe between C-G and G-C base pairs. Finally Chapter-8 discusses an overview of all the results.

CHAPTER - 2

MATERIALS AND METHODS

MATERIALS

The deoxyoligonucleotides 5' - 3' d-CpG, d-GpC, d-GpCpGpC and d-CpCpGpG as their sodium salts were purchased from M/s Pharmacia Fine Chemicals Sweden/P.L. Biochemicals, U.S.A. and were stored at -20°C. The tetrapeptide Lys-Trp-Gly-Lys OtBu 3 Ac.OH was purchased from M/s Bachem Feinchemikalien AG Switzerland and tripeptides Lys-Tyr-Lys. (2H COOH) and Lys-Phe-Lys.acetate were purchased from M/s Serva Feinbiochemica GmbH W. Germany and were used without purification. High purity D_2O 99.96% was obtained from M/s Merck Sharp & Dohme Ltd. Canada and M/s ICN Biomedicals Inc. Cambridge.

SAMPLE PREPARATION

The oligonucleotide and peptide samples were dissolved in known volume of solvent D_2O and the concentration were determined from measurements of absorbance using Beckman DU-6 UV-visible spectrophotometer. From the stock solution of known concentration thus prepared, a fixed volume was taken and diluted to another known volume to get the required concentration. The values of extinction coefficient used were

Oligonucleotides (value given by M/s P.L.Biochemicals)	Molar extinction coefficient)
d-CpG, d-GpC d-GpCpGpC, and d-CpCpGpG	at 260 nm, $9.02 \times 10^{3} M^{-1} Cm^{-1}$

Oligopeptides	Molar extinction coefficient
Lys-Trp-Gly-Lys-OtBu	at 280 nm, 5700 M^{-1} cm^{-1}
Lys-Tyr-Lys	at 280 nm, 1300 M^{-1} cm^{-1}
Lys-Phe-Lys	at 250 nm 200 M^{-1} cm^{-1}

The pH of the sample solution was adjusted to 7.0 by addition of small amounts of HCl or NaOH. The solutions were exchanged with D_2O (99.96%) at room temperature followed by lyophilization done twice. Finally the sample was redissolved in 500 µl D_2O (99.96%) and transferred to a 5 mm NMR tube (Wilmad). Typically 1 µl of 0.1 M solution of sodium 2, 2 Dimethyl - 2 - silapentane-5-Sulfonate (DSS) was added to the sample solution as internal standard. Further, EDTA was added most often in a concentration one tenth of the oligonucleotide or peptide concentration. For binding studies 2:1 mixture of the oligonucleotide (strand conc.) and oligopeptides were used so that peptide/duplex concentration ratio was one. The various concentrations used for oligonucleotides/peptide/mixture solutions for different systems are as follows:

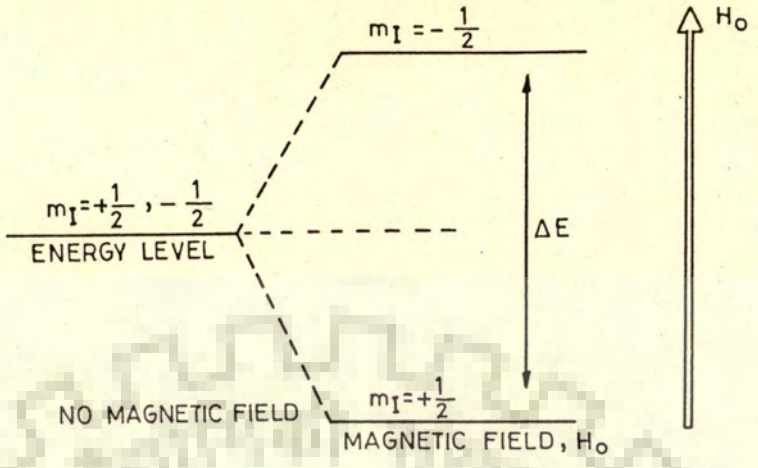
For d-CpG, 22 mM (strand concentration stated always hereafter) solution with 0.25 mM EDTA was used. For taking the spectra of oligopeptides alone 2.5 - 5.5 mM concentration sample solutions were used. A mixture of 5.0 mM d-CpG and 2.5 mM Trp or Phe containing oligopeptide was used whereas in case of Tyr containing peptide 11.0 mM d-CpG was mixed with 5.5 mM Lys-Tyr-Lys. For the tetranucleotides d-GpCpGpC

and d-CpCpGpG 19.8 mM concentration sample was used and then to 9.9 mM oligonucleotide solution, 4.95 mM Lys-Tyr-Lys was added. For measurements on d-GpC, 22 mM concentration sample was used and then to one half part of it i.e.11 mM GpC, 5.5 mM Lys-Trp-Gly-Lys OtBu was added and to other half 5.5 mM Lys-Tyr-Lys was added.

NMR METHODS

A. Theory

The phenomenon of nuclear magnetic resonance was discovered in 1946 by Felix Bloch and Edward Purcell [12]. It was later found that the exact frequency depends on the details of molecular environment which led chemists to exploit the technique for probing the structure of molecules. This analytical technique [147, 148, 68, 19] is based on the absorption of radiofrequency electromagnetic radiation by atomic nuclei, mainly protons but also other nuclei having an odd number of protons, in substances placed in a very strong magnetic field. A nucleus having magnetic moment I in a magnetic field H_o takes up (2I + 1) orientations which are characteristic of energies dependent on the magnitudes of magnetic moment μ and magnetic field H_o . For protons having I = 1/2, it is restricted to two possible orientations m_I = + 1/2 and -1/2 as shown in the figure on next page.
The two orientations correspond to two energy states and it is possible to induce transitions between them. The frequency



of the electromagnetic radiation which will effect such a transition is given by the equation

$$\nu = \frac{\gamma H_o}{2\pi}$$

where ν is the resonance frequency, γ is the gyromagnetic ratio of the nucleus. A normal NMR spectrum is characterised by several parameters the positions, widths, intensities and multiplicities of its lines. The chemical shift, δ defines the location of a NMR line along radio frequency axis and reflects the electronic environment around each nucleus site and is generally given relative to that of a reference as

$$\delta\,(\text{ppm}) = \frac{\nu\,\text{ref} - \nu\,\text{sample}}{\nu\,\text{ref}} \times 10^6$$

It is reported in parts per million (ppm) of the resonance frequency of the nucleus. The magnitude of δ is determined by the diamagnetic effects of circulating electron currents,

paramagnetic effects from neighbouring atoms and interatomic currents. In a system D-H...A the proton resonance is shifted downfield as the acceptor atom A pulls electrons and effects the diamagnetic currents associated with D-H so that deshielding occurs. In a benzene ring the six π electron current generate magnetic field as shown in Fig. 2.1a.

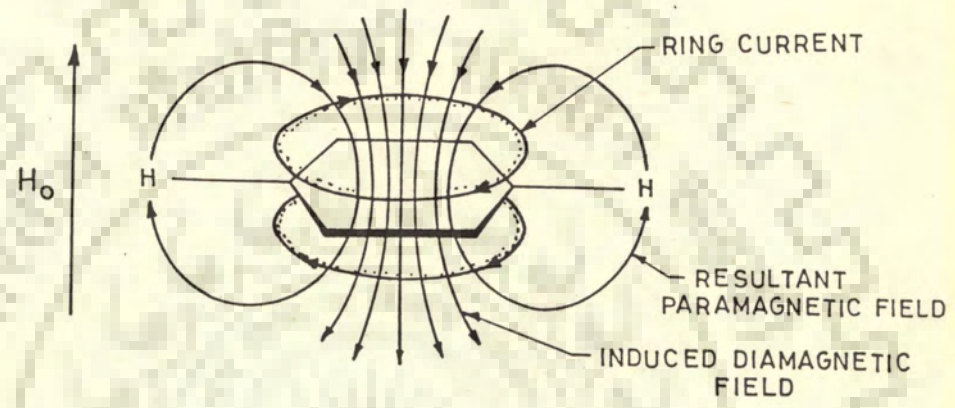


Fig. 2.1 (a) The deshielding of aromatic protons due to a ring current effect.

The ring current magnetic field acts in the same direction as the main field H_o so that the pendant protons are deshielded. However, the reverse is true for a proton placed above and below the ring and such a proton experiences upfield shift. These are particularly of significance for bases, base pairs, aromatic rings of peptides Trp, Tyr, Phe etc. Pullman and coworkers [47] have calculated these upfield shifts from the magnetic anisotropy effects and are shown in Fig.2.1 b-c for any proton placed above or below the plane of ring at a distance of 3.4 Å. Experimentally observed chemical shifts on stacking of base/peptide over a base/base pair could thus be used to find overlap geometry of aromatic rings.

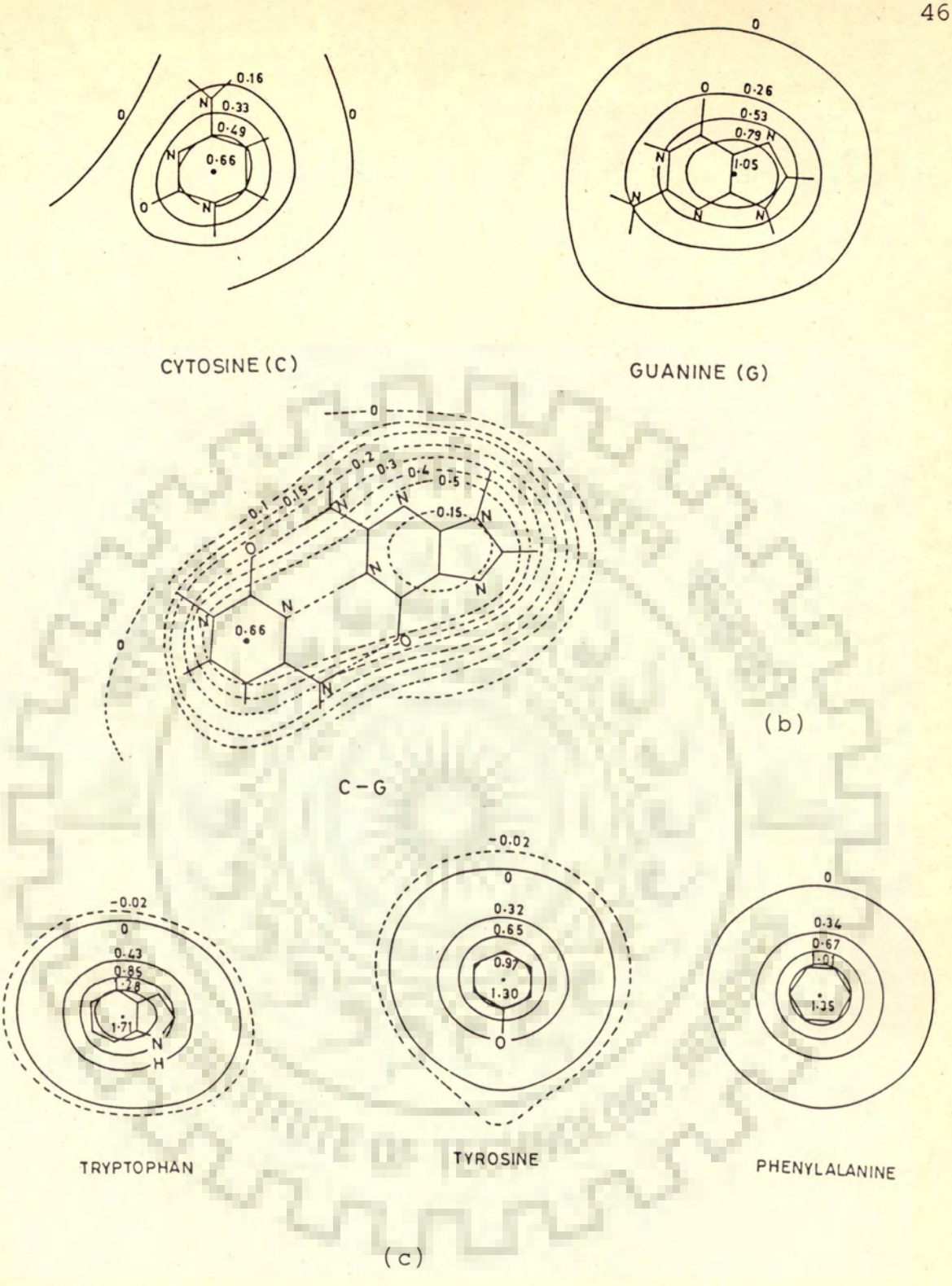
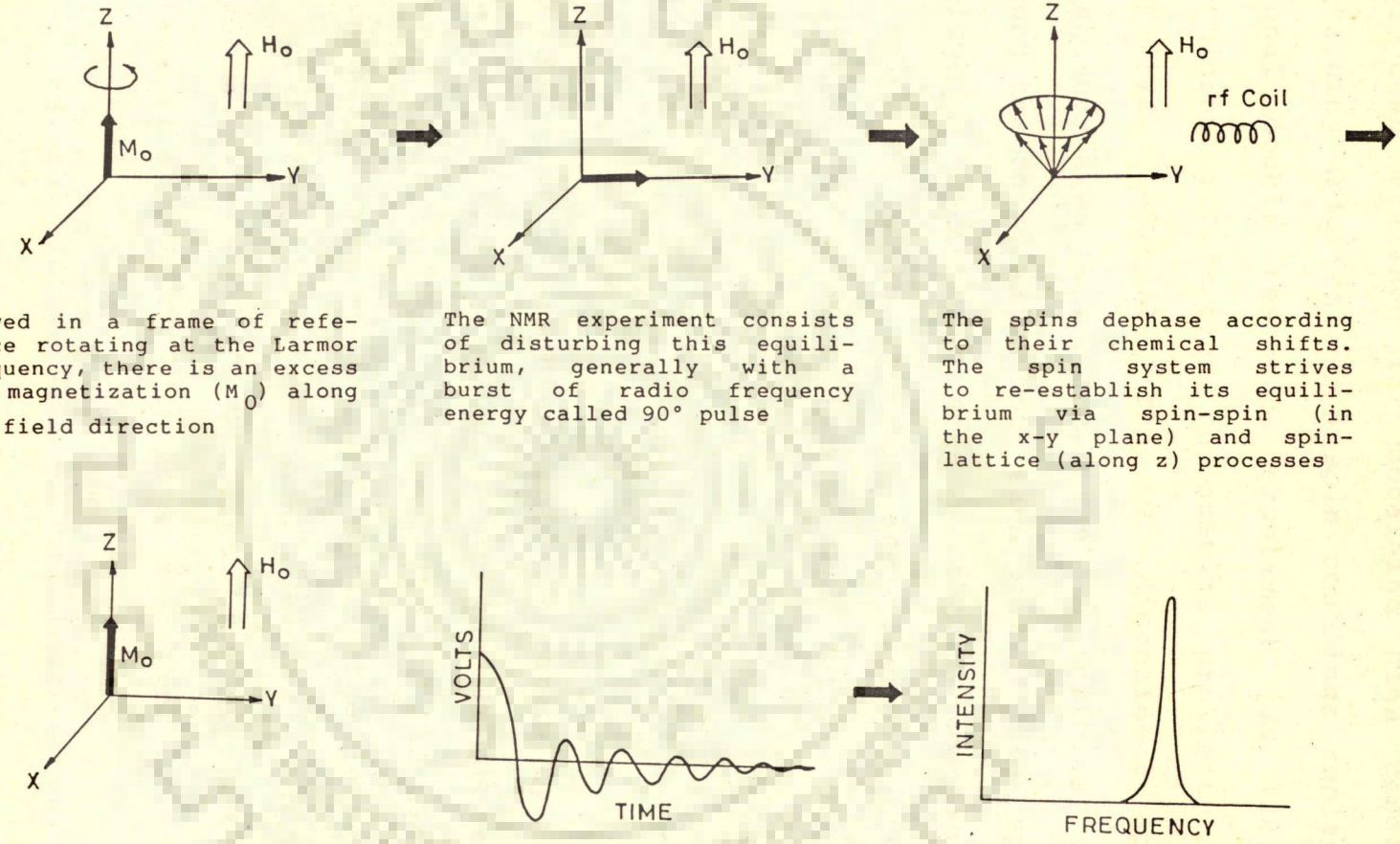


Fig. 2.1 **(b)** Intermolecular isoshielding curves (upfield shifts, $\Delta\delta$ in ppm) due to the sum of the contributions of the ring current and of the magnetic anisotropy in a plane at 3.4Å from the molecular surface of Cytosine, Guanine and C-G base pair **(c)** Tryptophan, Tyrosine and Phenylalanine.

Spin-spin splitting arises from the effects of neighbouring nuclei on the field felt by the nucleus of interest. The interaction is mediated by bonding electrons and depends on the distance between nuclei, the type of chemical bond,the bond angle and nuclear spin. The three bond spin-spin coupling constant ^{3}J (H-C-C - H) is a function of dihedral angle and therefore, used in finding conformation of various molecules and changes in conformation due to interactions.

In order to carry out an NMR experiment one needs a static magnetic field H_o for the alignment of the nuclear spins on RF field to stimulate absorption and a detector to record the resonance signals. In an FT-NMR spectrometer the sample is excited not by application of a continuous monochromatic RF field but by a short RF pulse which contains all the frequencies in the range of possible chemical shifts in a given nucleus. The magnetization M_o of the sample in the direction of the field H_o is reduced and a transverse magnetization builds up. The total magnetization is thus tilted through an angle which depends on the energy of pulses. Fig. 2.2 shows schematic illustration of the principle of pulsed FT spectroscopy.

The **two dimensional nuclear magnetic resonance (2D NMR) experiments** using the two pulse sequence (90° - t_1 - 90° - acquisition) was first proposed by Jeener [69]. The concept was further analysed in detail by Ernst 'S research group [8]. The technique



Viewed in a frame of reference rotating at the Larmor frequency, there is an excess net magnetization (M_0) along the field direction

The NMR experiment consists of disturbing this equilibrium, generally with a burst of radio frequency energy called 90° pulse

The spins dephase according to their chemical shifts. The spin system strives to re-establish its equilibrium via spin-spin (in the x-y plane) and spin-lattice (along z) processes

Equilibrium magnetization is eventually re-established and cycle can be repeated

Meanwhile, the induced voltage is monitored in the x-y plane by a resonant radio frequency coil. This signal is the free induction decay (FID)

The FID is Fourier transformed to give the desired frequency domain spectrum

Fig. 2.2 Schematic illustration of the principle of pulsed FT spectroscopy.

experiment uses the pulse sequence

$$90° - t_1 - 90° - t_2$$

where t_1 and t_2 are the evolution and detection period respectively [8]. In the pulse scheme of Fig.2.4a initially the spin system is allowed to relax to equilibrium during preparation period, at the end of which a 90° pulse converts the longitudinal magnetization into transverse magnetization. During the evolution period t_1 the magnetizations process at their resonant frequencies just as in a normal free induction decay. The second 90° pulse acts as a mixing pulse and transfers magnetization between J-coupled spins in the system. Some spins do not exchange magnetization and these spins give rise to peaks along the diagonal in the 2D-COSY spectrum. Both diagonal and off diagonal cross peaks have multiplet structures characteristic of the spins from which they originate. Thus a COSY spectra produces correlation map that displays connectivities by scalar spin-spin coupling and thus provide information on proximity of the nuclei along the chemical bonds [8, 96].

The correlation spectroscopy which reflects J-coupling correlation alone is not sufficient for obtaining a complete spectral assignment unless the approximate three-dimensional conformation of the molecule is known. In other words information about the proximity of atoms in space should also be available. This is achieved in a **two-dimensional nuclear overhauser**

enhancement spectrum (2D-NOE) which reflects dipolar coupling correlations between protons and indicates which pairs of protons are close by in space. Nuclear overhauser effect is the fractional change in intensity of signal in the spectrum by cross relaxation of one NMR line, when another resonance is perturbed. A detailed discussion on NOE is available in literature [99, 10]. The 1D experiments for measurements of NOE build up rates in macromolecules have been developed. But their practical use is limited since they require long accumulation times and because of poor selectivity for preirradiation of individual resonance lines in crowded region of 1D ^{1}H NMR spectra. 2D-NOESY is a more powerful and more efficient method for studies of selective NOE's between neighbouring protons in the spatial structure of biological macromolecules.

The method involves the application of a sequence of three successive nonselective 90° pulse [86, 70] Fig.2.4b The first 90° pulse creates transverse magnetization. During the evolution period t_1 the various magnetization components precess with their characteristic precession frequency in the X-Y plane of the rotating frame, are thus frequency labelled. After the second 90° pulse cross-relaxation leads to incoherent magnetization exchange during the mixing time τ_m. The signal is recorded immediately after the third pulse as a function of t_2; t_3 is a fixed relaxation delay to enable the system to reach equilibrium after each recording. The experiment

constants since the components of cross peaks have antiphase character and tend to cancel each other when the resolution is not enough to resolve the J separation between them. As a result no cross peaks are expected for very low J values that ~0-3 Hz, weak cross peaks may be seen for values of 4-7 Hz and very strong peaks will be seen for values of 8-10 Hz. The schematic spectra expected for the various sugar geometries are shown in Fig 2.6 and summarized in Table 2.2. These can be divided into five principle types:

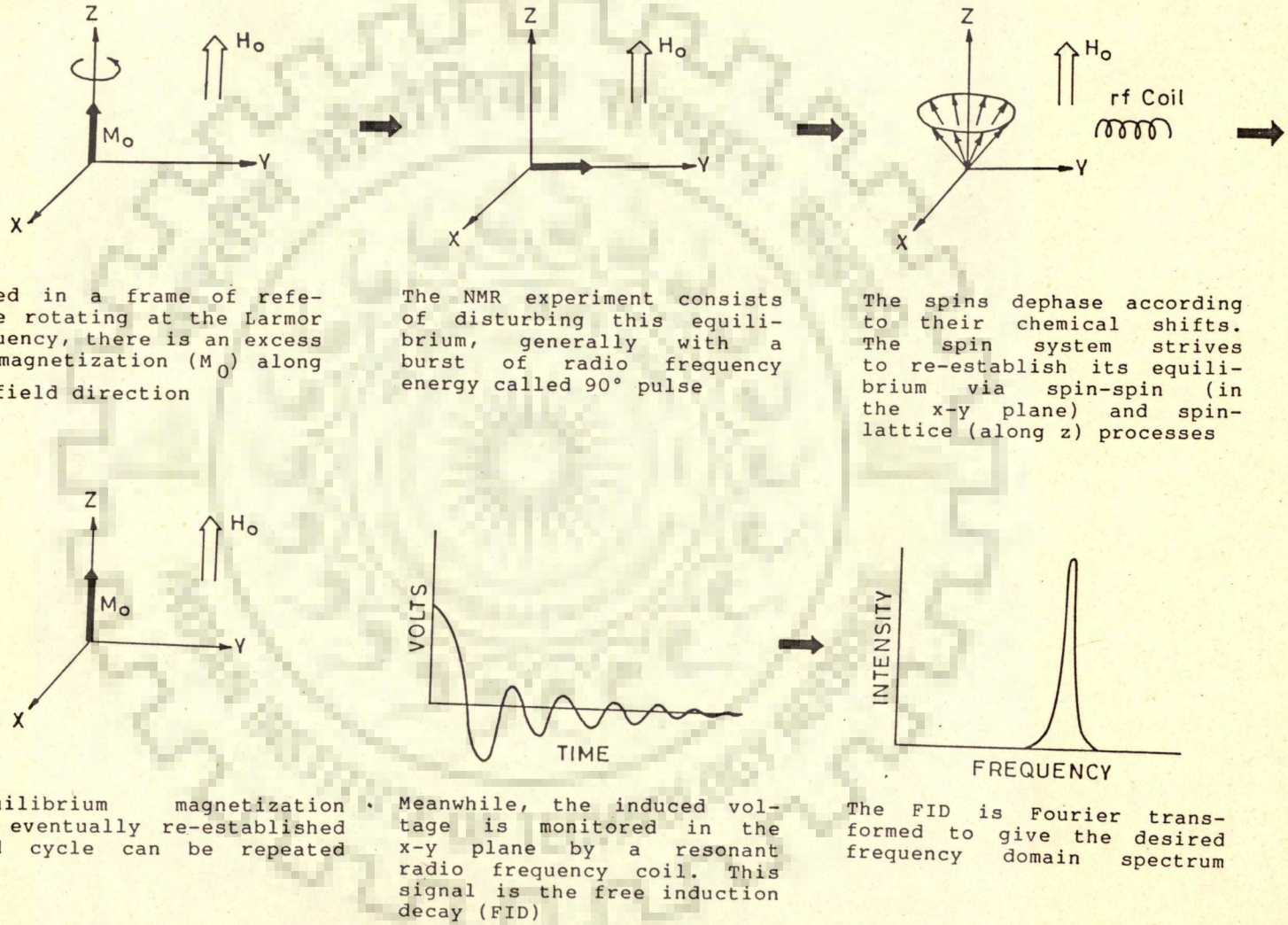
Type I — Sugar conformations C3'— endo, C2'— exo and C4'— exo for which COSY spectra is characterised by absence of H1' - H2' cross peak.

Type II — Sugar conformation O1'— endo (O4'— endo) in which there is no connectivity between H2" and H3'.

Type III — Sugar conformation C2'— endo, C3'— exo and C1'— exo for which there are no cross peaks between H2" - H3' and between H3' - H4'.

TYPE IV — Sugar conformation O1'— exo (O4'— exo) and C4'— endo which are identified by the absence of H1' - H2', H2" - H3' and H3' - H4' connectivities.

Type V — Sugar conformation C1'— endo for which the connectivities H1 - H2' and H3' - H4' are absent.



Viewed in a frame of reference rotating at the Larmor frequency, there is an excess net magnetization (M_0) along the field direction

The NMR experiment consists of disturbing this equilibrium, generally with a burst of radio frequency energy called 90° pulse

The spins dephase according to their chemical shifts. The spin system strives to re-establish its equilibrium via spin-spin (in the x-y plane) and spin-lattice (along z) processes

Equilibrium magnetization is eventually re-established and cycle can be repeated

Meanwhile, the induced voltage is monitored in the x-y plane by a resonant radio frequency coil. This signal is the free induction decay (FID)

The FID is Fourier transformed to give the desired frequency domain spectrum

Fig. 2.2 Schematic illustration of the principle of pulsed FT spectroscopy.

has been utilised for obtaining high resolution [95] for obtaining information on spin couplings replacing the conventional double resonance experiment [8, 139, 74, 145] and obtaining information on spatial proximity of protons via NOE [74, 145, 73]. The two-dimensional spectrum has two frequency axes (ω_1, ω_2) and the intensities are represented along the third axis. Thus each peak in a 2D spectrum occupies volume as against an area in the 1D spectrum. The two frequencies in a 2D spectrum are generated by Fourier transformation of a time domain data matrix of two independent time variables, t_1, t_2.

$$S(t_1, t_2) \rightarrow S(\omega_1, \omega_2)$$

The two time variables are generated by segmentation of the conventional time axis of the FT-NMR experiments. The period t_1 is called evolution period, t_2 is termed as the detection period and it is only during the latter that the data is collected. Fig.2.3a shows general subdivision of the time axis of a 2D-FT NMR expt. and Fig.2.3b shows the pulse scheme of the 2D-NMR (Jeener's experiment).

Several 2D pulse sequences have since been discovered on the principle of transfer of magnetization from one nucleus to another. The two dimensional correlation spectroscopy (2D-COSY) has found widespread use for the identification of coupling partners in networks of coupled spins [69, 8, 96, 9]. The

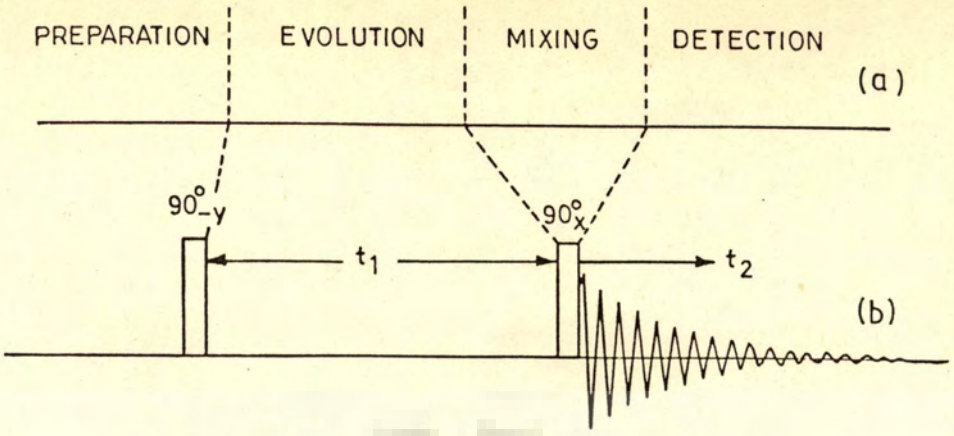


Fig. 2.3 **(a)** General subdivision of the time axis into four phases: Preparation, Evolution, Mixing and Detection periods, of which Evolution and Detection periods contain the two time Variables t_1 and t_2.
(b) Pulse scheme of the Basic 2D-NMR (Jeener's experiment).

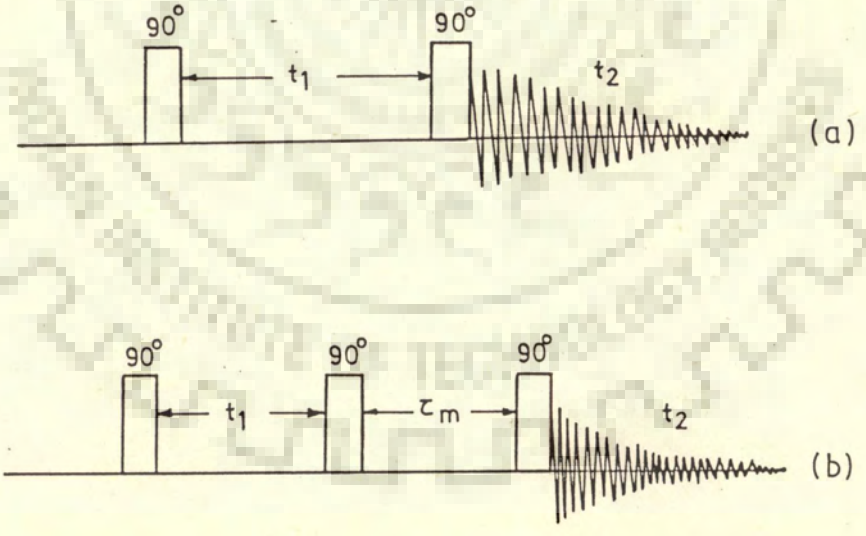


Fig. 2.4 Fundamental Pulse sequence for **(a)** 2D-COSY NMR and **(b)** 2D-NOESY NMR technique.

experiment uses the pulse sequence

$$90° - t_1 - 90° - t_2$$

where t_1 and t_2 are the evolution and detection period respectively [8]. In the pulse scheme of Fig.2.4a initially the spin system is allowed to relax to equilibrium during preparation period, at the end of which a 90° pulse converts the longitudinal magnetization into transverse magnetization. During the evolution period t_1 the magnetizations process at their resonant frequencies just as in a normal free induction decay. The second 90° pulse acts as a mixing pulse and transfers magnetization between J-coupled spins in the system. Some spins do not exchange magnetization and these spins give rise to peaks along the diagonal in the 2D-COSY spectrum. Both diagonal and off diagonal cross peaks have multiplet structures characteristic of the spins from which they originate. Thus a COSY spectra produces correlation map that displays connectivities by scalar spin-spin coupling and thus provide information on proximity of the nuclei along the chemical bonds [8, 96].

The correlation spectroscopy which reflects J-coupling correlation alone is not sufficient for obtaining a complete spectral assignment unless the approximate three-dimensional conformation of the molecule is known. In other words information about the proximity of atoms in space should also be available. This is achieved in a two-dimensional nuclear overhauser

enhancement spectrum (2D-NOE) which reflects dipolar coupling correlations between protons and indicates which pairs of protons are close by in space. Nuclear overhauser effect is the fractional change in intensity of signal in the spectrum by cross relaxation of one NMR line, when another resonance is perturbed. A detailed discussion on NOE is available in literature [99, 10]. The 1D experiments for measurements of NOE build up rates in macromolecules have been developed. But their practical use is limited since they require long accumulation times and because of poor selectivity for preirradiation of individual resonance lines in crowded region of 1D ^{1}H NMR spectra. 2D-NOESY is a more powerful and more efficient method for studies of selective NOE's between neighbouring protons in the spatial structure of biological macromolecules.

The method involves the application of a sequence of three successive nonselective 90° pulse [86, 70] Fig.2.4b The first 90° pulse creates transverse magnetization. During the evolution period t_1 the various magnetization components precess with their characteristic precession frequency in the X-Y plane of the rotating frame, are thus frequency labelled. After the second 90° pulse cross-relaxation leads to incoherent magnetization exchange during the mixing time τ_m. The signal is recorded immediately after the third pulse as a function of t_2; t_3 is a fixed relaxation delay to enable the system to reach equilibrium after each recording. The experiment

is repeated for a set of equidistant t_1 values adequate signal-to-noise ratio, n transients are accumulated for each value of t_1. Two-dimensional fourier transformation of the data matrix $S(t_1, t_2)$ then produces the desired frequency domain spectrum $S(\omega_1, \omega_2)$. Peaks corresponding to the one-dimensional spectrum appear on the diagonal and NOE connectivities between individual lines are manifested by pairs of cross peaks in symmetrical locations with respect to the diagonal peak [54]. The measurements thus yield interatomic distances (separation of two protons by a distance of say 3 to 4Å) and hence lead to determination of molecular structure and conformation.

B. Experimental

All proton NMR experiments were carried out at DST sponsored National FT-NMR Facility at Tata Institute of Fundamental Research, Bombay and were recorded on a 500 MHz high resolution Bruker AM 500 FT-NMR spectrometer equipped with Aspect 3000 computer The typical parameters for 1D-NMR experiments at different temperature were pulse duration 12 μsec (60° pulse), no. of data points 8K - 16K. Spectral width 5000 Hz, no. of scans 60-1000 and a resolution of 0.15 Hz/points to 1.22 Hz/point,. Receiver gain value was optimised in every experiment to get the best signal-to-noise ratio in limits of good sensitivity. All 2D-COSY and -NOESY experiments were carried out at room temperature, 297K. The typical

NMR parameters were 1024-2048 data points along the t_2 axis and 256-512 data points along t_1 axis; no. of scans 60 to 200, pulse width of 15 μ sec. and sweep width 5000 Hz, a resolution of 1.00 Hz/point in both ω_1, ω_2 dimensions, Relaxation delay of 1.00 sec. and a mixing time (τ_m) of 800 m sec for NOESY experiments.

C. Strategies used for conformational analysis

The sugar conformation, sequential resonance assignment, helix sense, glycosidic bond rotation, duplex formation and inter-proton distances are determined from the 2D-COSY and 2D-NOESY experiments.

The COSY spectra gives information about the J-coupling constants, the values of which are fixed for a deoxyribose depending upon its puckering mode. The five ^3J values namely H1' - H2', H1'-H2" H2'-H3', H2"-H3' and H3'-H4' are related to the relevant H-C$\overset{\wedge}{-}$C-H dihedral angle (Ø) by the equation.

$$^3_J = 10.5 \cos^2\theta - 1.2 \cos\theta$$

The values of J-coupling constants for various dihedral angles and their relationship with pseudorotation parameter P have been found by Altona and Sundaralingam [3]. These are plotted in Fig.2.5 and tabulated in Table 2.1 Under usual conditions for recording COSY spectra the intensities of the cross peaks depend directly on the magnitude of the coupling

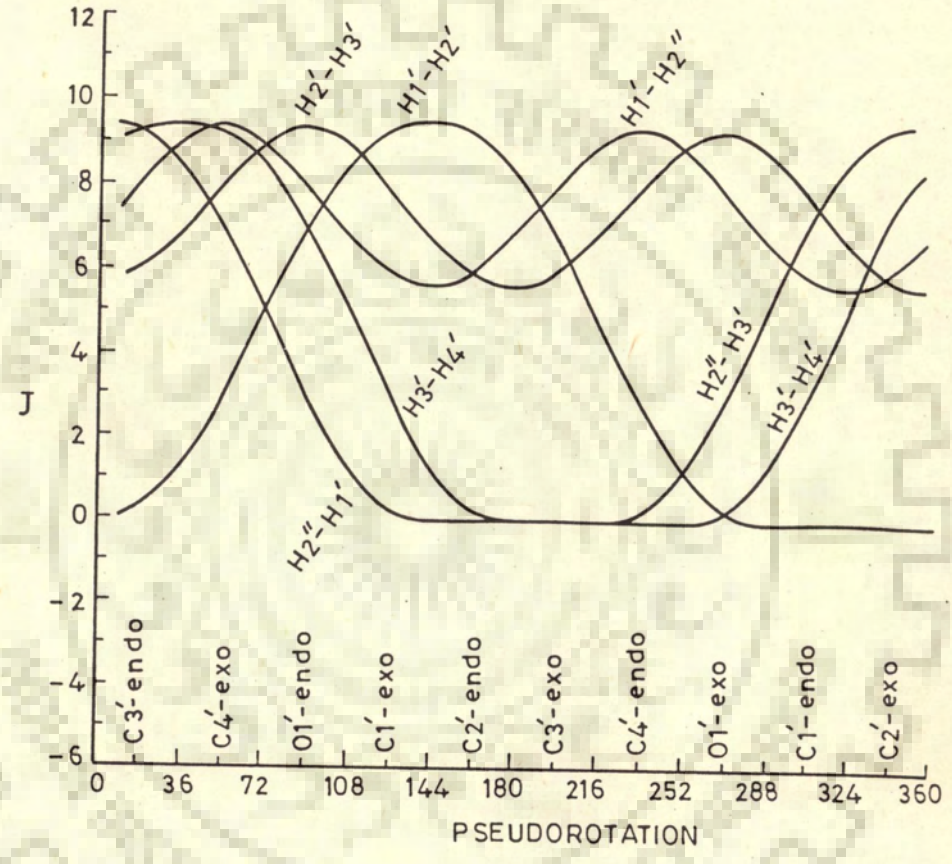


Fig. 2.5 Variation of the vicinal coupling constants in the deoxyribose ring as a function of the ring geometry.

Table 2.1 : Calculated ^{3}J-Coupling Constants (in Hz) in a Pentose Ring for various Envelope Conformations (3)

Sugar Pucker (value of P in parentheses)	$J_{1'-2'}$	$J_{1'-2''}$	$J_{2'-3'}$	$J_{2''-3'}$	$J_{3'-4'}$
1) C3' -endo (18°)	0.1(0.28)	7.4(8)	4.9(6)	11.0(9.28)	10.3(9.14)
2) C4' -exo (54°)	2.4 (2.7)	9.1 (9.14)	7.0 (7.7)	9.0 (7.14)	10.3 (9.28)
3) O1' -endo or O4' -endo (90°)	(6.7)	(8.0)	(9.28)	(3.14)	(7.14)
4) C1' -exo (126°)	10.1 (9.14)	5.5 (6)	7.6 (8)	0.1 (0.28)	2.9 (3)
5) C2' -endo (162°)	10.1 (9.14)	5.5 (6)	5.5 (6)	0.1 (0)	0.2 (0.28)
6) C3' -exo (198°)	7.9 (7.0)	7.6 (8)	5.5 (6)	0.1 (0)	0 (0)
7) C4' -endo (234°)	(2.7)	(9.14)	(6.7)	(0.14)	(0.0)
8) O1' -exo or O4' -exo (270°)	(0.28)	(8.14)	(9.28)	(2.56)	(0.14)
9) C1' -endo (306°)	(0.0)	(6.14)	(8.14)	(6.56)	(2.7)
10) C2' -exo (342°)	0.2 (0.0)	5.0 (6.0)	4.9 (6.14)	11.0 (9.14)	8.2 (6.7)

Note : The values given in brackets have been calculated from Figure 2.5.

constants since the components of cross peaks have antiphase character and tend to cancel each other when the resolution is not enough to resolve the J separation between them. As a result no cross peaks are expected for very low J values that ~0-3 Hz, weak cross peaks may be seen for values of 4-7 Hz and very strong peaks will be seen for values of 8-10 Hz. The schematic spectra expected for the various sugar geometries are shown in Fig 2.6 and summarized in Table 2.2. These can be divided into five principle types:

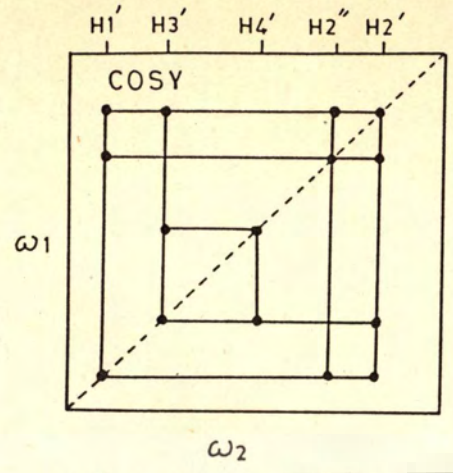
Type I – Sugar conformations C3'— endo, C2'— exo and C4'— exo for which COSY spectra is characterised by absence of H1' - H2' cross peak.

Type II–Sugar conformation O1'— endo (O4'— endo) in which there is no connectivity between H2" and H3'.

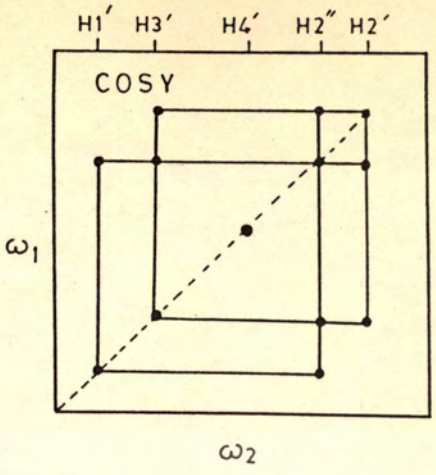
Type III–Sugar conformation C2'— endo, C3'— exo and C1'— exo for which there are no cross peaks between H2" - H3' and between H3' - H4'.

TYPE IV–Sugar conformation O1'— exo (O4'— exo) and C4'— endo which are identified by the absence of H1' - H2', H2" - H3' and H3' - H4' connectivities.

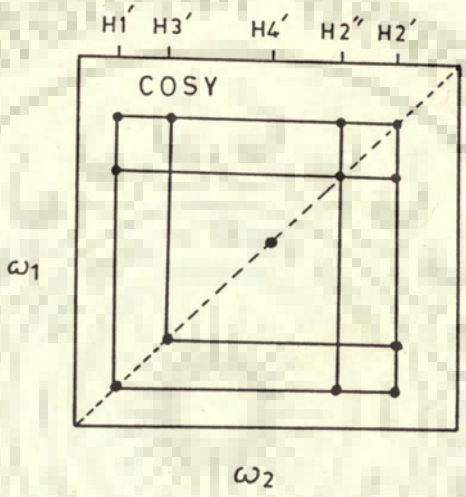
Type V— Sugar conformation C1'— endo for which the connectivities H1 - H2' and H3' - H4' are absent.



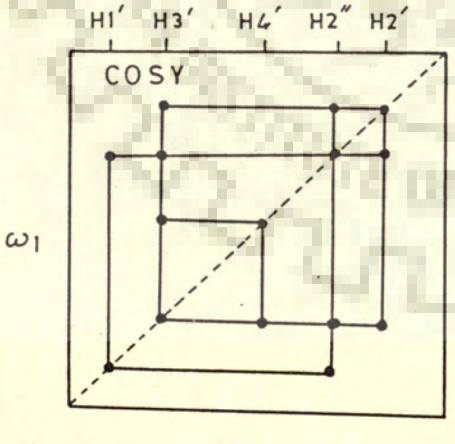
O1′-endo or O4′-endo



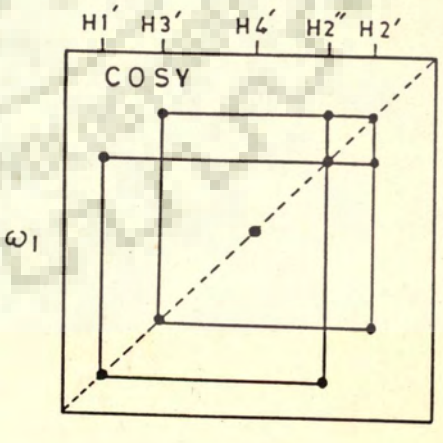
C1′-endo



C2′-endo or C1′-exo or C3′-exo



C3′-endo or C2′-exo or C4′-exo



C4′-endo or O1′/O4′-exo

Fig. 2.6 Schematic COSY spectra for the various sugar geometries.

Table 2.2 : Various Interproton Connectivities for a Deoxyribose Sugar ring, in a COSY Spectra, under low resolution conditions, as a function of P.

Type of Sugar Pucker	P (in°)	Cross Peaks H2'-H2"	H1'-H2"	H1'-H2'	H2'-H3'	H2"-H3'	H3'-H4'
1. C3'-endo, C2'-exo and C4'-exo (Type I)	-18 to 54	+	+	-	+	+	+
2. O1'-endo or O4'-endo (Type II)	90	+	+	+	+	-	+
3. C1'-exo, C2'-endo and C3'-exo (Type III)	126 to 198	+	+	+	+	-	-
4. C4'-endo and O1'-exo or O4'-exo (Type IV)	234 to 270	+	+	-	+	-	-
5. C1'-endo (Type V)	306	+	+	-	+	+	-

Note : + indicates the presence of connectivity between the two protons.

- indicates the absence of connectivity between the two protons.

Additional connectivities H4' - H5', H4' - H5" and H5' - H5" also depicted in COSY spectra do not help in assigning the sugar pucker. Thus having identified H1', H2', H2", H3' and H4' peaks in any NMR spectra, from the cross peaks in COSY spectra it is possible to determine sugar conformation for each nucleotide residue.

The 2D-NOESY spectra provides a complete sequential resonances assignment [66]. For right-handed structures with sugars in C3'-endo/C2'-endo/O1'-endo pucker conformations for all values of glycosidic angle in the anti domain, a convenient strategy for sequential assignment is

$(GH8/CH6)_n$ ..... $(H2', H2", H1')_n$

intra-residue NOE

$(GH8/CH6)_n$ ..... $(H2", H1')_{n-1}$

inter-residue NOE

where n stands for nth residue in 5' - 3' oligonucleotide sequence. Schematically these may be shown as in Fig. 2.7a.

For left-handed z-DNA conformations, however, the strategies are different and are as follows:

(GH8) ... H1' intra-residue NOE

(CH6) ... H5",H2' - do -

$(GH8)_n$... $(H5")_{n-1}$ inter-residue NOE

$(CH6)_n$... $(H2')_{n-1}(H3')_{n-1}$ - do -

$(CH5)_n$... $(H2')_{n-1}$ - do -

Schematically this may be shown as in Fig.2.7b.

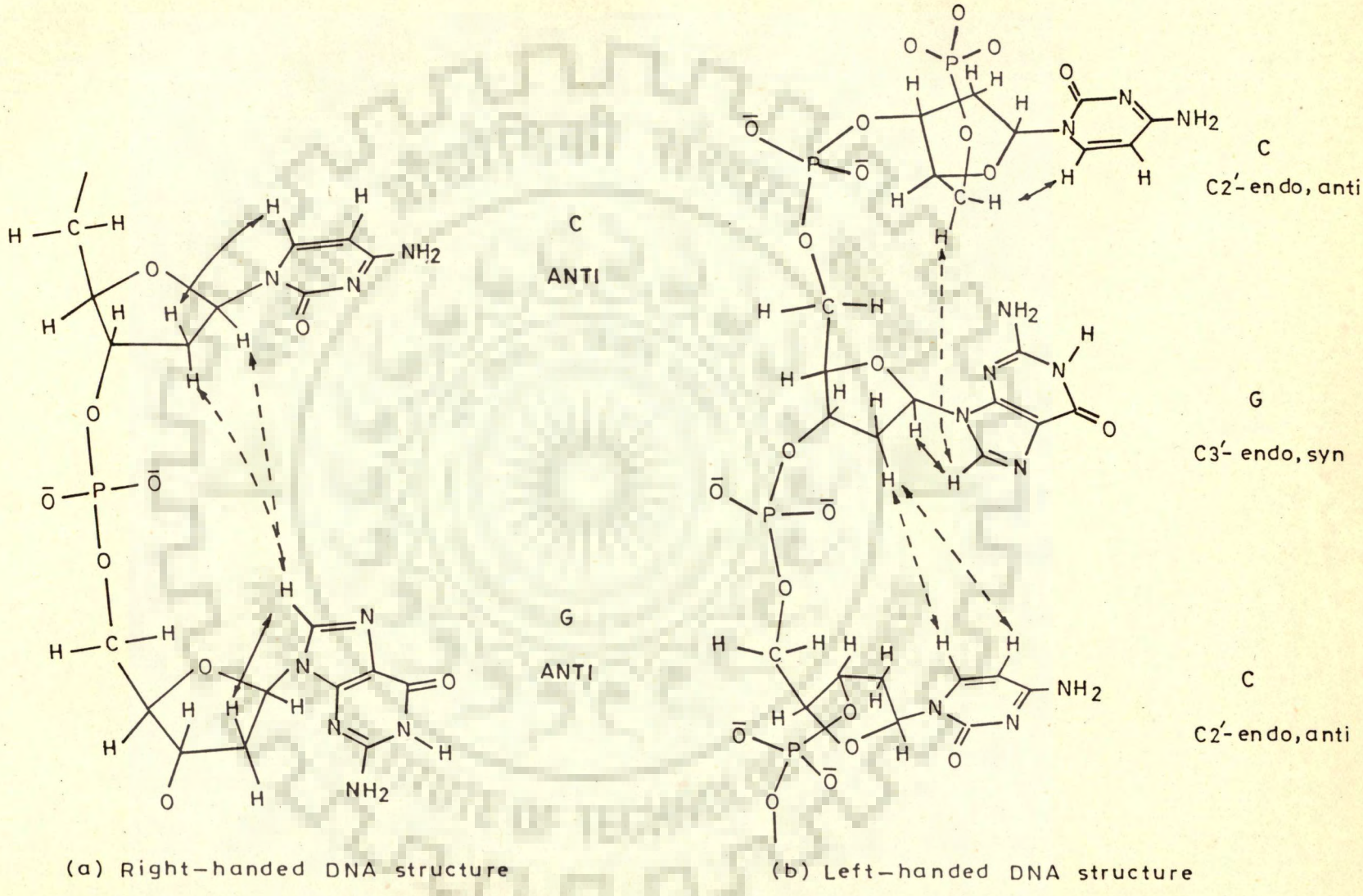


Fig. 2.7 Strategies of sequential resonance assignment from NOESY spectra of oligonucleotides for **(a)** right-handed helices **(b)** left-handed, z-type structure.

The glycosidic bond conformation, that is, the orientation of the base relative to the sugar is described by the torsional angle χ_{CN}. A discrimination between anti, high anti, and syn conformations is made on the basis of strength of NOE's between base protons GH8/CH6 and corresponding H1', H2', H2" sugar protons. For a Syn conformation a strong NOE is expected to be observed between H8/H6 and H1' protons. At the same time NOE's from base to H2' and H2" protons should be relatively weak and have different intensities.

For an anti conformation, NOE's from base H8/H6 to H2', H2" protons of the same nucleotide should have different intensities. Also the proton corresponding to the weaker NOE should show a strong NOE to the base proton of the next nucleotide. For the high-anti conformation the H2" protons should show strong NOE's to the base protons of the same nucleotide and the following nucleotide on 3'-end.

Interproton distances have been obtained from the distance calculations existing in literature [147]. Mazeau et al [90] have determined the interproton distances within furanose ring as a function of pseudorotation of sugar cycle for $d(Br^5CG)_3$ (Fig. 8). Govil et al [52] have obtained interproton distances as a function of pseudorotation and glycosidic angle in the form of contour plots (Fig. 2.9 a-d) for base to H2', H2", H3' and H4" sugar protons. Using dinucleotide d-CpG as a representative example, for the dihedral angles $\psi=36°$, $\psi'=83°$, $\Phi' = 180°, \omega = 314°, \omega' = 314°, \Phi = 214°$ and sugar conformation

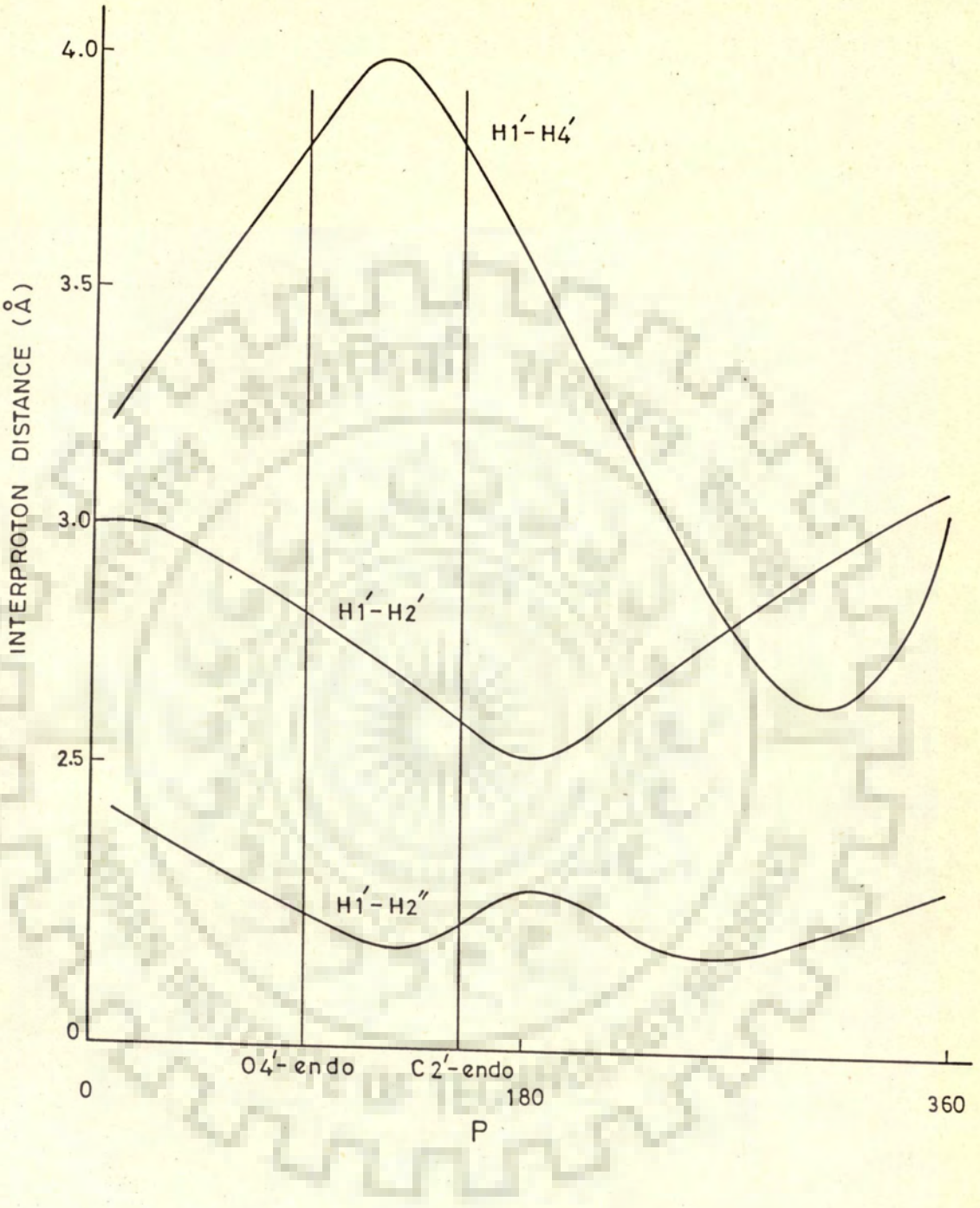


Fig. 2.8 Interproton distances as a function of pseudorotation of sugar cycle.

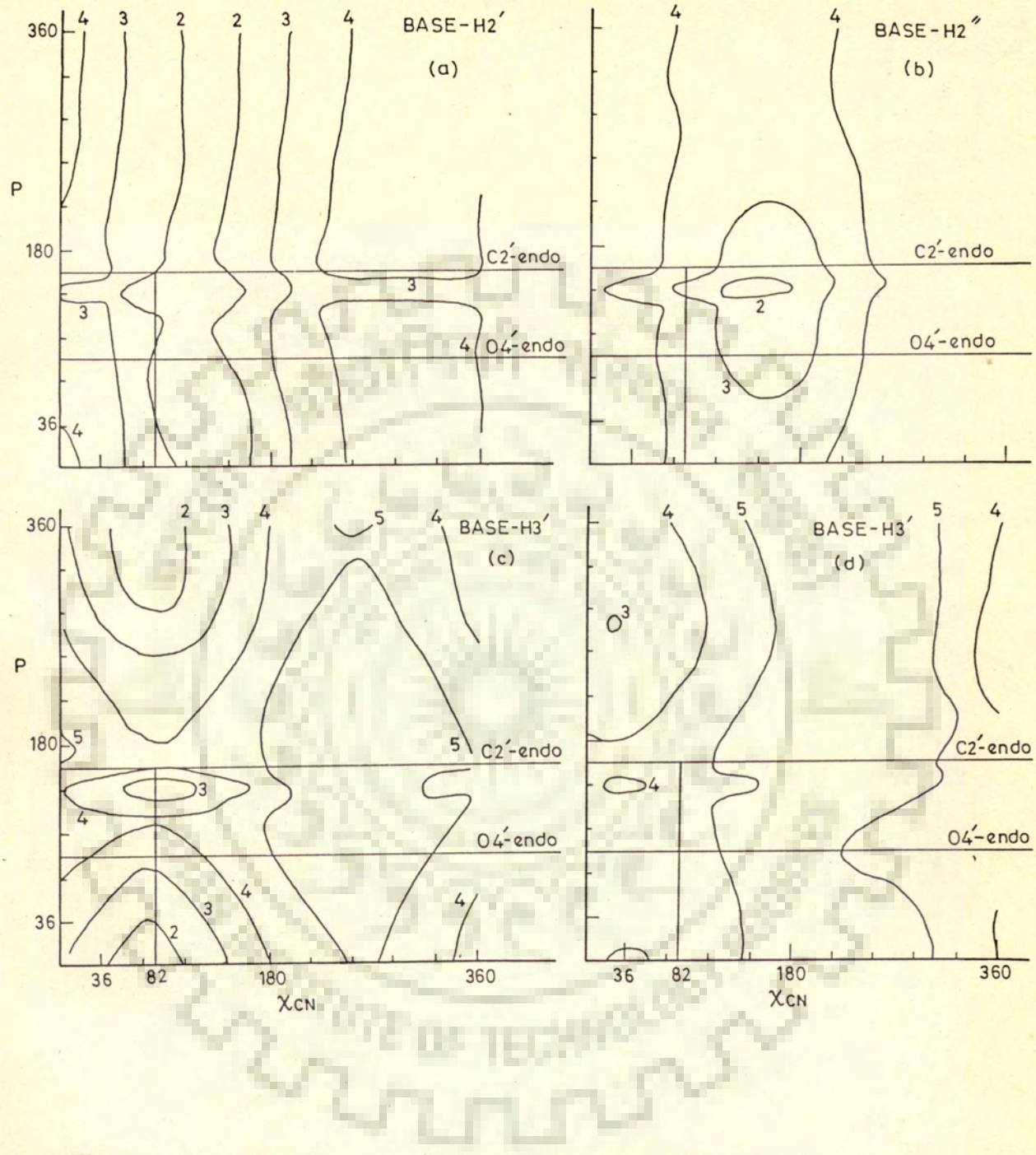


Fig. 2.9 Contour plots showing interproton distances as a function of pseudorotation, P and glycosidic bond angle, χ for Base H6/H8 to **(a)** H2' **(b)** H2" **(c)** H3' (d) H4' sugar protons.

as C3'—endo/C2'—endo for bases G and C, intra—residue H8/H6... H1', H2', H2" distances and inter—residue GH8 ... CH2" distances have been obtained (Fig. 2.10) as a function of glycosidic dihedral angle [66]. We have made use of these to compare the distances obtained from our model building studies.

MODEL BUILDING

Model building studies were done for d-GpG and its complexes with oligopeptides containing Trp and Tyr residues by using Dreiding skeletal stereo-models obtained from M/s Buchi Laboratorium Switzerland. The model units are made of small rods and tubes radiating from and connected to a point which represents the atomic nucleus. Rods and tubes correspond in their numbers and spatial arrangment to the sigma bonds characteristic of principle atoms C, O, N, etc. The atomic nuclei are represented by different colours e.g. carbon is denoted by black colour, oxygen by red, nitrogen by blue, sulphur by yellow and phosphorous by grey colour. Their lengths are proportional (0.4Å/cm or 2.5 cm/Å) to the inter-nuclear distance between a hydrogen and the represented principle atoms so that all unmarked free ends of rods and tubes signify centres of hydrogen atoms. More complicated models are constructed by pushing the rod of one unit into the tube of another.

Coordinates given by Arnott and Hukins [7] based on X-ray diffraction studies of the crystals of various families

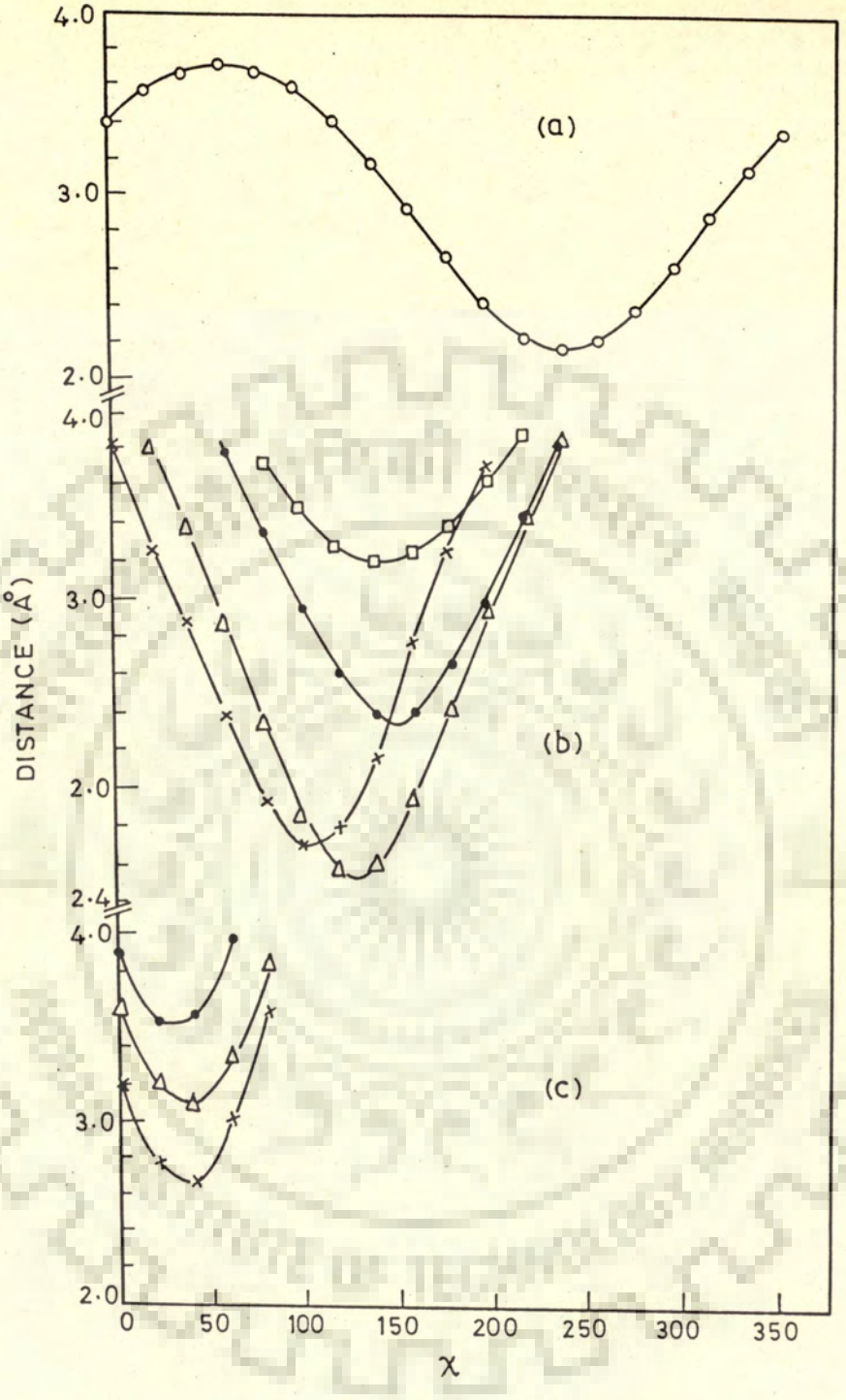


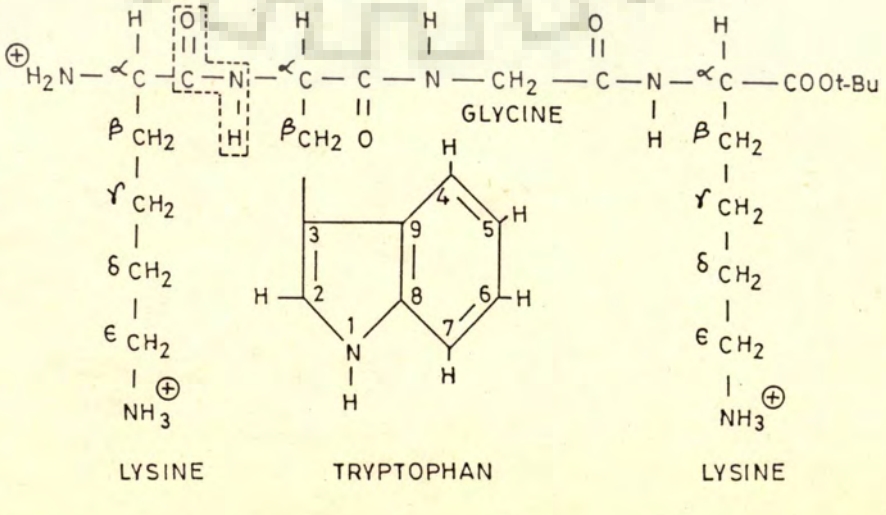
Fig. 2.10 Interproton distances as a function glycosidic dihedral angle: **(a)** intra-residue H8/H6----H1' **(b)** intra-residue H8/H6----H2', H2" for C3'-endo and C2'-endo geometries, --X--X--X-- is for H2' in C2'-endo pucker, --●--●--●-- is for H2" in C2'-endo pucker; ----□----□---- and ---△---△---△-- are the corresponding curves for C3'-endo geometry, **(c)** inter-residue distances GH8---CH2" for the two sugar geometries, ---●---●--- and ---△---△---△-- are for CH2' and CH2" respectively for C2'-endo geometry. For C3'-endo pucker only CH2" proton shows a shrot distance with GH8.

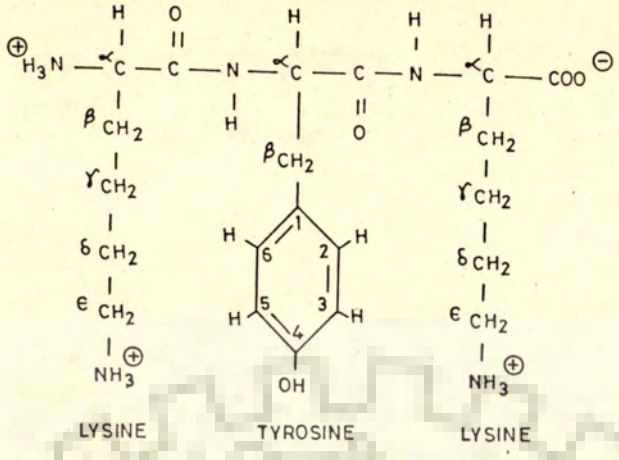
of DNA were used to construct the model of self complementary duplex of d-CpG. The values of the backbone angles were

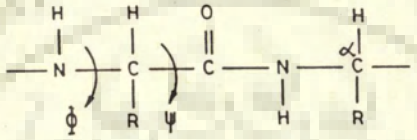
$$O3'-P\overset{\omega}{\rightleftharpoons}O5'\overset{\Phi}{\rightleftharpoons}C5'\overset{\psi}{\rightleftharpoons}C4'\overset{\psi'}{\rightleftharpoons}C3'\overset{\Phi'}{\rightleftharpoons}O3'\overset{\omega'}{\rightleftharpoons}P$$

$$=46^\circ \quad =-147^\circ \quad =36^\circ \quad =157^\circ \quad =155^\circ \quad =-96^\circ$$

For C2'—endo sugar conformation the value of pseudorotation, P = 162° and the torsional angles τ_0 and τ_1 are -21.5° and 36.0°, respectively. For O1'—endo sugar the value of P is 90° so that τ_2 and τ_3 are 0° and -22.3°, respectively. The glycosidic bond rotation χ_{CN} is defined as angle O4'-C1' $\rightleftharpoons$ N9-C8 in purines and as angle O1' - C1' $\rightleftharpoons$ N1 - C6 in pyrimidines. For χ_{CN} as an anti conformation a value of -98° or 262° was used. This corresponds to 82° according to the definition used by Govil et al [51] that is O1' - C1' - N9 - C8 for purines and O1' - C1' - N1' - C6 for pyrimidines. Models of Lys-Trp-Gly-Lys OtBu and Lys-Tyr-Lys were made by taking Φ, ψ - angles as -58° and -46°.







The models of complexes of d-CpG with Trp and Tyr containing oligopeptides were made by intercalating Trp or Tyr exactly in the middle of the base pairs, the distance between which was increased from 3.4 to 6.8Å. This had resulted in a change in backbone angles. The position of Trp or Tyr with respect to the base pairs above and below was selected to be such that with the resulting base pair overlap geometry the upfield shifts in Trp or Tyr ring protons was in order with that obtained from magnetic anisotropy curves of Pullman and coworkers [47] given in Figure 2.1.a-e.

THEORETICAL METHODS

We have estimated the conformational energy using classical potential function (CPF). Such an approach nas widely been used in the study of protein structure [119, 127], nucleic acid geometry [80, 102, 149, 50] and drug nucleic acid interaction [24, 100]. In this approximation

the total potential energy, V, is assumed to be

$$V = V_a + V_r + V_{el} + V_\theta$$

where,

V_a = Van der waal's attraction energy
V_r = repulsion energy,
V_{el} = electrostatic energy, and
V_θ = torsional energy.

The interactions due to stretching and bending of covalent bonds are neglected as the vibrational constants are considerably higher than the torsional barriers.

Van der waal's attractive energy also known as London's dispersion energy arises from the interaction between fluctuating, transient dipoles on the two atoms concerned and is of the form $\sum_{i<j} - A/R_{ij}^6$, where i and j are non bonded atoms and A_{ij} depends on the polarisabilities of atoms i and j. An expression for A_{ij} has been derived as :

$$A_{ij} = \frac{\frac{3}{2}\, e\hbar/\sqrt{M}.\ \alpha_i.\ \alpha_j}{\sqrt{\frac{\alpha_i}{N_i}} + \sqrt{\frac{\alpha_i}{N_j}}}$$

where α_i, α_j are the polarisabilities on atoms i and j and N_i, N_j are the effective number of polarisable electrons on atoms i and j [131].

The repulsion energy expression

$$V_r = \sum_{i<j} \frac{B_{ij}}{R_{ij}^{12}}$$

is an empirical one where B_{ij} is calculated by imposing a condition that the minimum in total potential energy occurs when the separation between pairs of atom is equal to the sum of Van der waal's radii R_o for a given pair of atoms i and j.

The electrostatic energy expression

$$V_{el} = \sum_{i<j} \frac{q_i q_j}{\varepsilon R_{ij}}$$

where q's are the partial charges on atoms arising due to the differences in electronegativity of atoms participating in covalent bonds and calculated by standard quantum mechanical approaches. The effective dielectric constant ε is taken as 4.0.

The torsional potential is characterised by the bond about which the rotation is considered and arises from the exchange interactions between the orbitals of the bonded atoms. It is taken as

$$V(\theta) = V1\ (1 + \text{Cos}\ 3(\theta)) + V2\ (1 + \text{Cos}\ 2(\theta))$$

where V_1, V_2 the barrier heights have been taken as 1.5.

The interaction energy between two molecules has been determined by treating the interaction as a perturbation to the isolated system and partitioned into physically meaningful energy terms. The details of method are reviewed by Claverie [125] and used by Govil and coworkers [76, 75, 77, 78]. The interaction energy is sum of electrostatic (E_{el}), polarisation (E_{pol}), dispersion (E_{disp}) and repulsion energies (E_{rep})

$$E_{int} = E_{el} + E_{pol} + E_{disp} + E_{rep}$$

The molecular charge distribution has been expressed as a set of atomic charges, q, and atomic dipoles , μ, obtained by the CNDO (Complete Neglect of Differential Overlap) method. For electrostatic interaction

$$E_{el} = E_{qq} + E_{q\mu} + E_{\mu\mu}$$

where,

E_{qq} is monopole-monopole interaction

$E_{q\mu}$ is monopole-dipole interaction

$E_{\mu\mu}$ is dipole-dipole interaction

The explicit expressions for these are:

$$E_{qq} = \sum_{i}^{(1)} \sum_{j}^{(2)} \frac{q_i\, q_j}{r_{ij}}$$

$$E_{q\mu} = \sum_{i} \sum_{j} q_i(\bar{\mu}_j \cdot \bar{R}_{ij}) \cdot |R_{ij}|^{-3}$$

$$E_{\mu\mu} = \sum_i \sum_j (\bar{\mu}_i, \bar{\mu}_j)|(R_{ij})|^{-3} - 3(\bar{\mu}_i.\bar{R}_{ij})(\bar{\mu}_j.\bar{R}_{ij})|R_{ij}|^{-5}$$

The **polarisation energy** of a binary complex is given as

$$E_{pol} = E_{pol(2\rightarrow1)} + E_{pol(1\rightarrow2)}$$

where $E_{pol(1\rightarrow2)}$ stands for energy due to polarisation of molecule 2 by 1. Further

$$E_{pol(2\rightarrow1)} = -\frac{1}{2}\sum_u^{(1)} \varepsilon_u^{(1)}.\ Au.\ \varepsilon_u^{(2)}$$

where $\varepsilon_u^{(2)}$ is electric field induced by molecule 2 at midpoint of bond u of molecule 1 and Au is the polarisability tensor of bond u of molecule 1.

The **dispersion** and **repulsion** terms considered together in Kitaigorskii type formulae is

$$E_{disp} + E_{rep} = \sum_i^{(1)} \sum_j^{(2)} E(i,j)$$

where,

$$E(I,j) = K_i K_j \left[\frac{-A}{Z_{ij}^6} + Be^{-\gamma Z_{ij}}\right]$$

$$Z_{ij} = \frac{r_{ij}}{R_{ij}}$$

$$R_{ij}^0 = [(2R_i^\omega)\ (2R_j^\omega)]^{\frac{1}{2}}$$

where R_i^{ω} and R_j^{ω} are Van der wall's radii of atoms i and j; A, B, γ are constants i.e. A = 0.214, B = 47000 and γ = 12.35; parameters K_i, K_j are taken from Cailet and Claverie [18]. The calculated interaction energies have been used to give insight into forces stabilising nucleic acid-protein associations.

CHAPTER - 3

d-CpG AND ITS BINDING TO OLIGOPEPTIDES CONTAINING Trp, Tyr AND Phe RESIDUES

500 MHz proton NMR spectra of deoxyoligonucleotide d-CpG, tripeptide Lys-Tyr-Lys, tetrapeptide Lys-Trp-Gly-Lys OtBu, tripeptide Lys-Phe-Lys and a mixture of d-CpG with these peptides as a function of temperature in the range 277-355 K were studied. Homonuclear two—dimensional chemical shift correlated spectroscopy (COSY) and nuclear overhauser enhancement spectroscopy (NOESY) spectra of these samples were taken to find their conformation.

d-CpG

Figure 3.1 shows a typical spectra of d-CpG in D_2O as a function of temperature. Nonexchangeable base protons and sugar protons are easily identified from the chemical shift values [51,147]. The assignment of specific sugar protons to G or C base has been made on the basis of NOESY spectra. The observed chemical shifts (Table 3.1) have been plotted in Figure 3.2. The change in chemical shift $\Delta\delta$ due to an increase in temperature in CH2' proton is large being about 0.18 ppm whereas that in GH4', GH8 protons is low, being about 0.02 ppm. These changes indicate that at lower temperatures d-CpG exists as double—helix which separates into

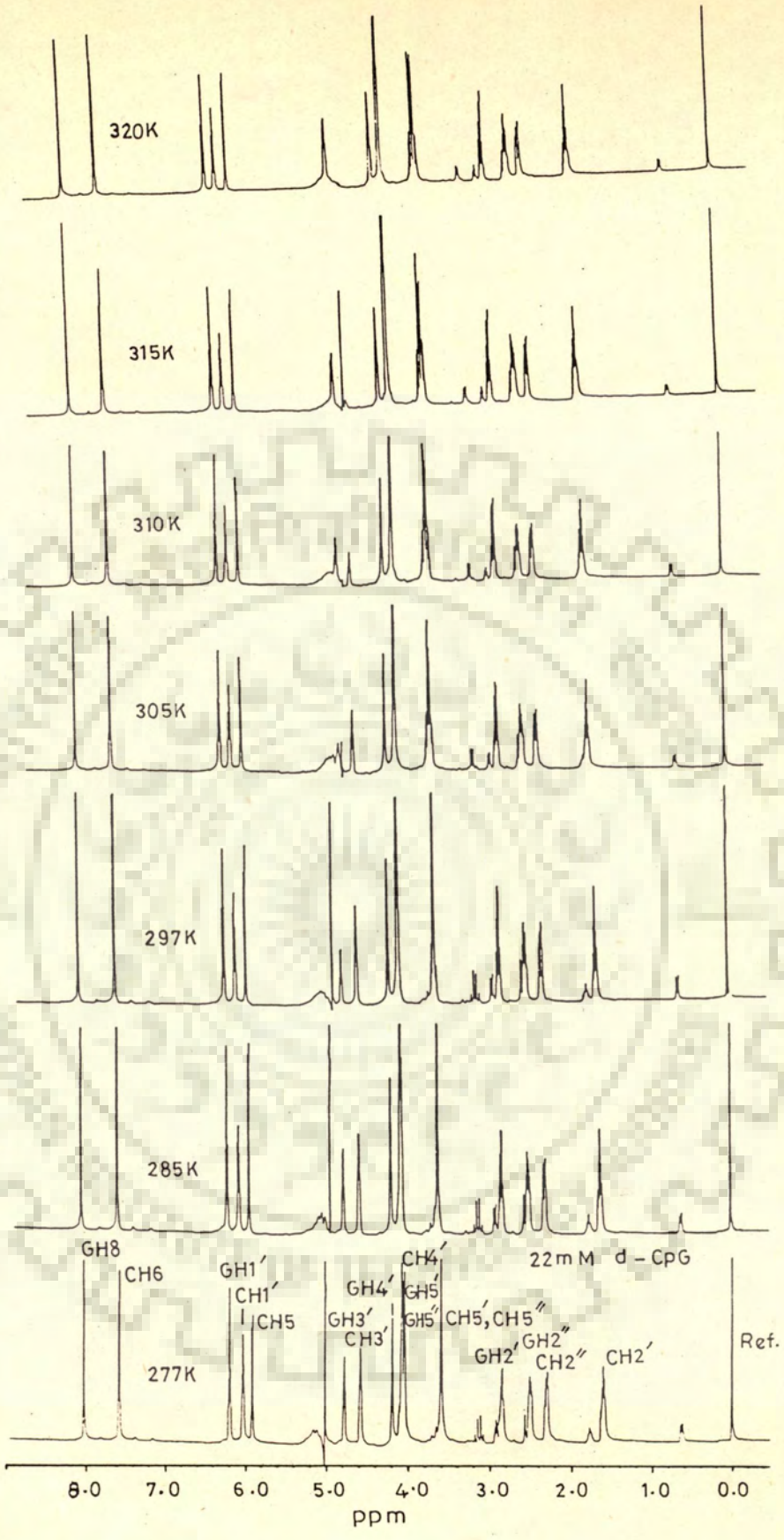


Fig. 3.1 500 MHz proton NMR spectra of 22 mM (Strand concentration) deoxydinucleotide d-CpG solution in H_2O (pH=7.1) containing d-CpG 0.25 mM EDTA at various indicated temperatures. Ref.DSS.

Table 3.1 : Chemical shift values of various protons at 500 MHz for 22 mM d-CpG in D_2O (pH=7.1) in temperature range 277K-320K

Temp. (K)	GH8	CH6	CH5	GH1'	CH1'	GH2'	GH2"	CH2'	CH2"
277	8.008	7.562	5.903	6.185	6.015	2.843	2.501	1.598	2.296
285	8.025	7.570	5.926	6.212	6.052	2.832	2.507	1.620	2.305
297	8.032	7.578	5.944	6.216	6.080	2.821	2.499	1.664	2.322
305	8.028	7.585	5.952	6.225	6.093	2.817	2.523	1.706	2.338
310	8.027	7.589	5.957	6.229	6.101	2.808	2.519	1.730	2.347
315	8.025	7.593	5.960	6.233	6.105	2.821	2.516	1.752	2.339
320	8.020	7.597	5.563	6.235	6.108	2.797	2.526	1.774	2.349

Temp. (K)	GH3'	CH3'	GH4'	CH4'	GH5'/H5"	CH5'/H5"
277	4.763	4.571	4.183	4.069	4.033	3.594
285	4.759	4.573	4.186	4.072	4.068	3.622
297	4.806	4.574	4.186	4.068	4.068	3.640
305	4.754	4.579	4.186	4.068	4.068	3.656
310	4.749	4.584	4.186	4.069	4.069	3.668
315	4.733	4.620	4.187	4.069	4.069	3.666
320	4.735	4.735	4.186	4.069	4.069	3.661

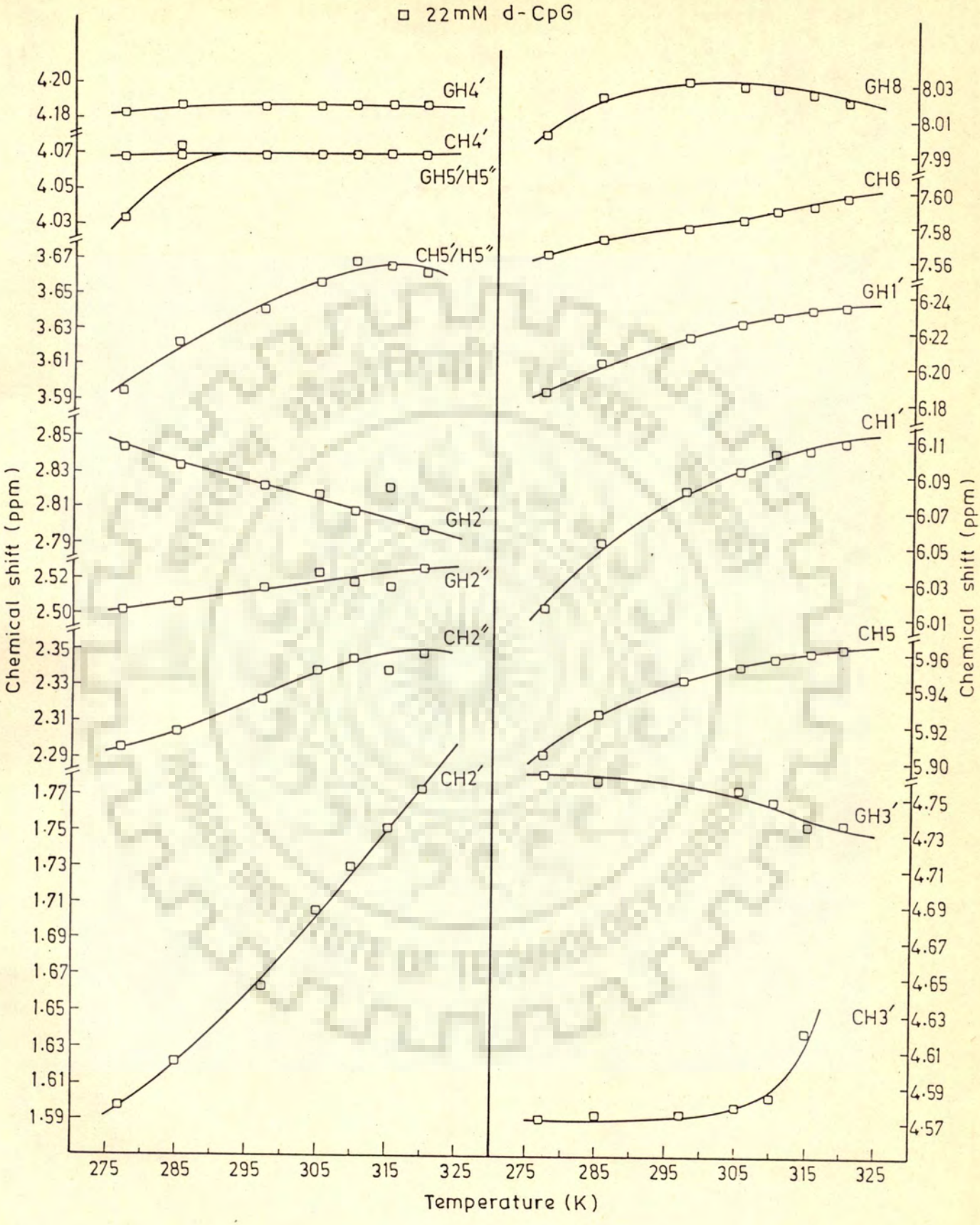


Fig. 3.2 Chemical shifts of base and sugar protons of 22 mM d-CpG as a function of temperature, from spectra in Fig. 3.1.

strands as the temperature is raised. The transition is rather broad probably because of less perfect double helical geomtery of a small dinucleotide. $T_{\frac{1}{2}}$ value defined by the midpoint of transition ranges between 295 to 297 K.

Figures 3.3a , 3b and 3.4 are the results of COSY and NOESY experiments performed at 297 K for the same sample solutions as in Figure 3.1. Figures 3.5a and 3.5b are the expansions of specific regions of the results of NOESY experiment on d-CpG. Assignments have been carried out in steps by using strategies discussed in Chapter two by taking into account spin—spin coupling and dipole-dipole interactions. Two sets of spin-spin coupled sugar proton resonances are easily seen in Figure 3.3. Since within a nucleotide base protons GH8 or CH6 are closest to the corresponding H2' (about 2.1 Å) than H2" and H1' (both, about 3.6 Å), specific assignment of two sets of sugars to G and C nucleotides has been made using results of Figure 3.5a. The two H1' protons resonating at 6.216 and 6.080 ppm have been assigned to G and C nucleotides, respectively. It may be noted that GH2' resonates at a higher ppm than GH2" proton. The assignment of GH2' and GH2" protons is definitive and unambiguous because H1' is closer to H2" as compared to H2' in righthanded DNA structures and we see clearly NOE connectivity of GH8 with GH2' (not GH2") and GH1' with GH2" (not GH2'). This is also consistent with the shape of GH2' and GH2" spectral lines in 1D - NMR (Figure 3.1). The result that GH2' resonates

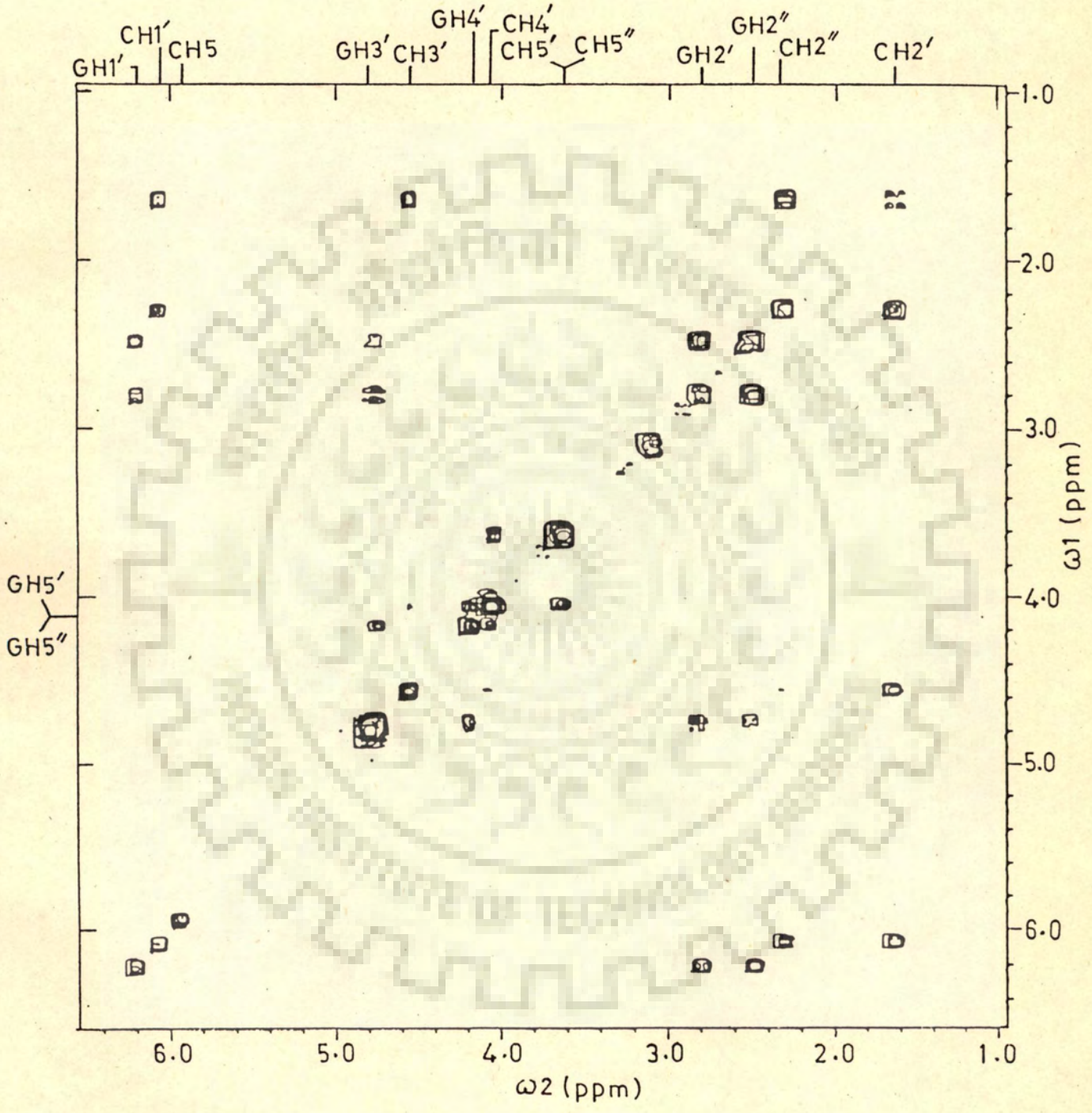


Fig. 3.3 (a) 500 MHz COSY spectrum of 22 mM d-CpG solution in D_2O containing 0.25 mM EDTA (pH=7.1) at 297 K in chemical shift range of 1-6.4 ppm.

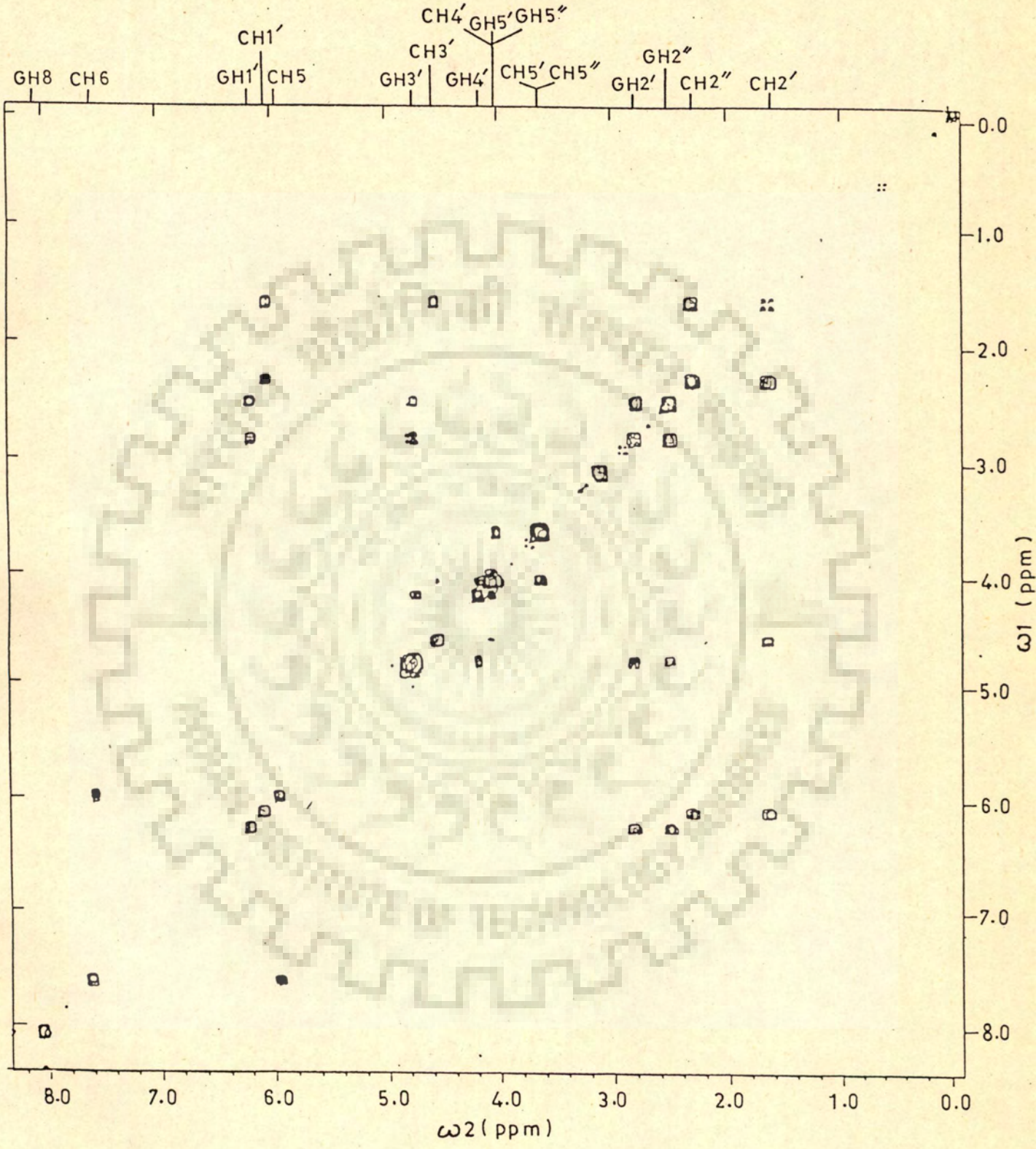


Fig. 3.3 (b) 500 MHz COSY spectrum of 22 mM d-CpG solution in D_2O containing 0.25 mM EDTA (pH=7.1) at 297 K in chemical shift range of 0-8 ppm.

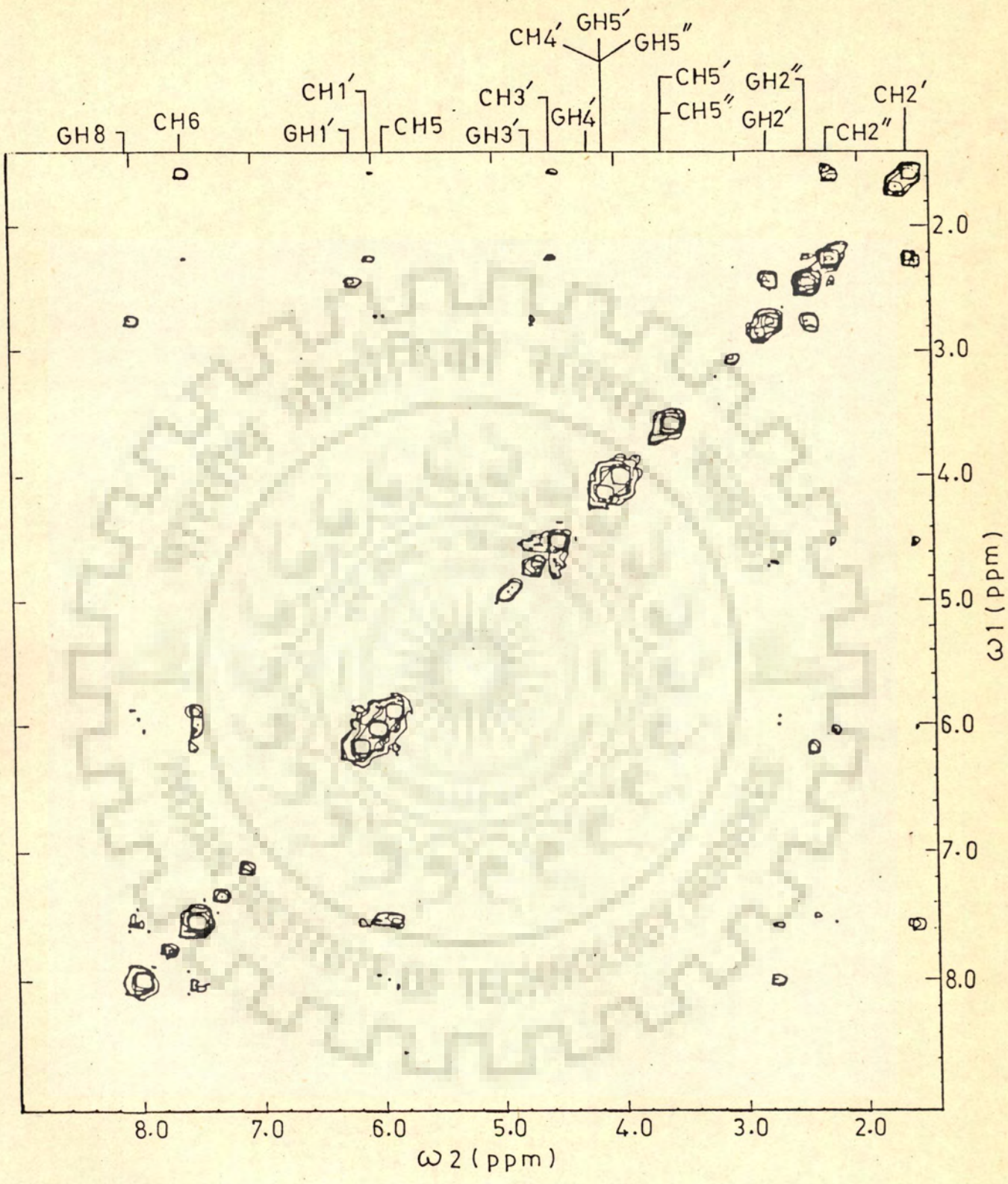
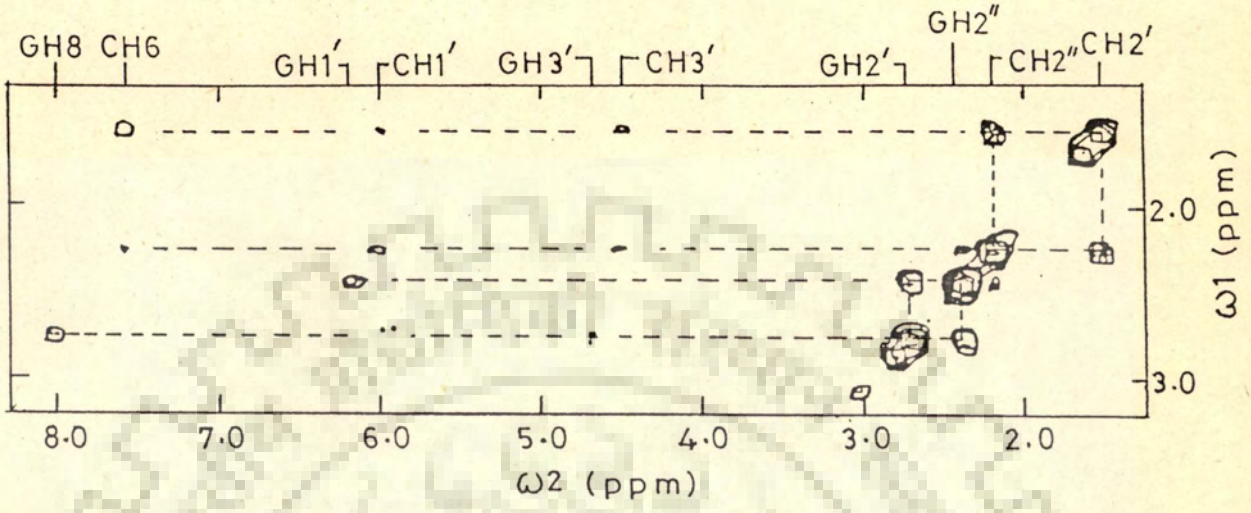
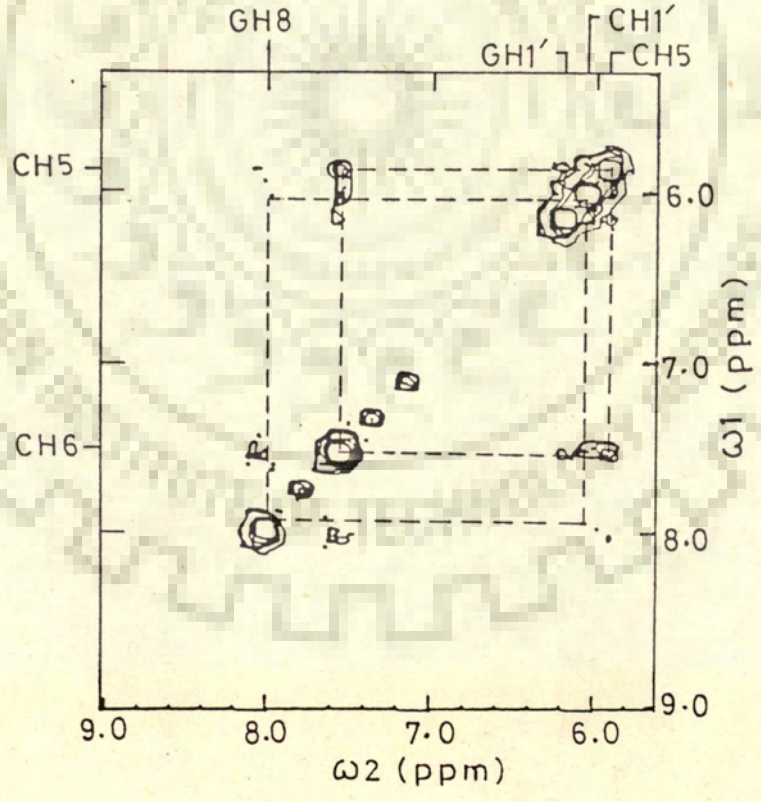


Fig. 3.4 500 MHz NOESY spectrum of 22 mM d-CpG solution in D_2O containing 0.25 mM EDTA (pH=7.1) at 297 K.



(a)



(b)

Fig. 3.5 Portions of NOE spectra of Fig. 3.4 expanded to show **(a)** base and H1' ... Sugar H2', H2", **(b)** base ... H1' and CH6...CH5 connectivities.

at a higher ppm than GH2" proton is however, contrary to that expected in deoxyoligonucleotides and to that found for cytosine nucleotide in d-CpG itself. This is possible due to the fact that d-CpG is particularly a small oligonucleotide and is not expected to form a perfect helix. Figure 3.5b also shows a weaker NOE connectivity of GH8 proton with GH1' as well as CH1' proton as expected on the basis of sequential assignment. A complete unambiguous assignment of various sugar protons is thus made.

The results on relative intensities of cross-peaks in COSY spectra, summarized and shown schematically in Figure 3.6, are as follows:

GH1' ... GH2' and GH2"	both equally strong	(1)
GH2' ... GH3'	strong	(2)
GH2" ... GH3'	strong	(3)
GH3' ... GH4'	strong	(4)
CH1' ... CH2' and CH2"	both equally strong	(5)
CH2' ... CH3'	strong	(6)
CH2" ... CH3'	weak (not seen clearly in figure)	(7)
CH3' ... CH4'	weak	(8)

Since cross peaks of H1' with both H2' and H2" are equally strong [(1) and (5)] and H3' shows J correlated connectivity with H4' [(4) and (8)] it is inferred that both G and C

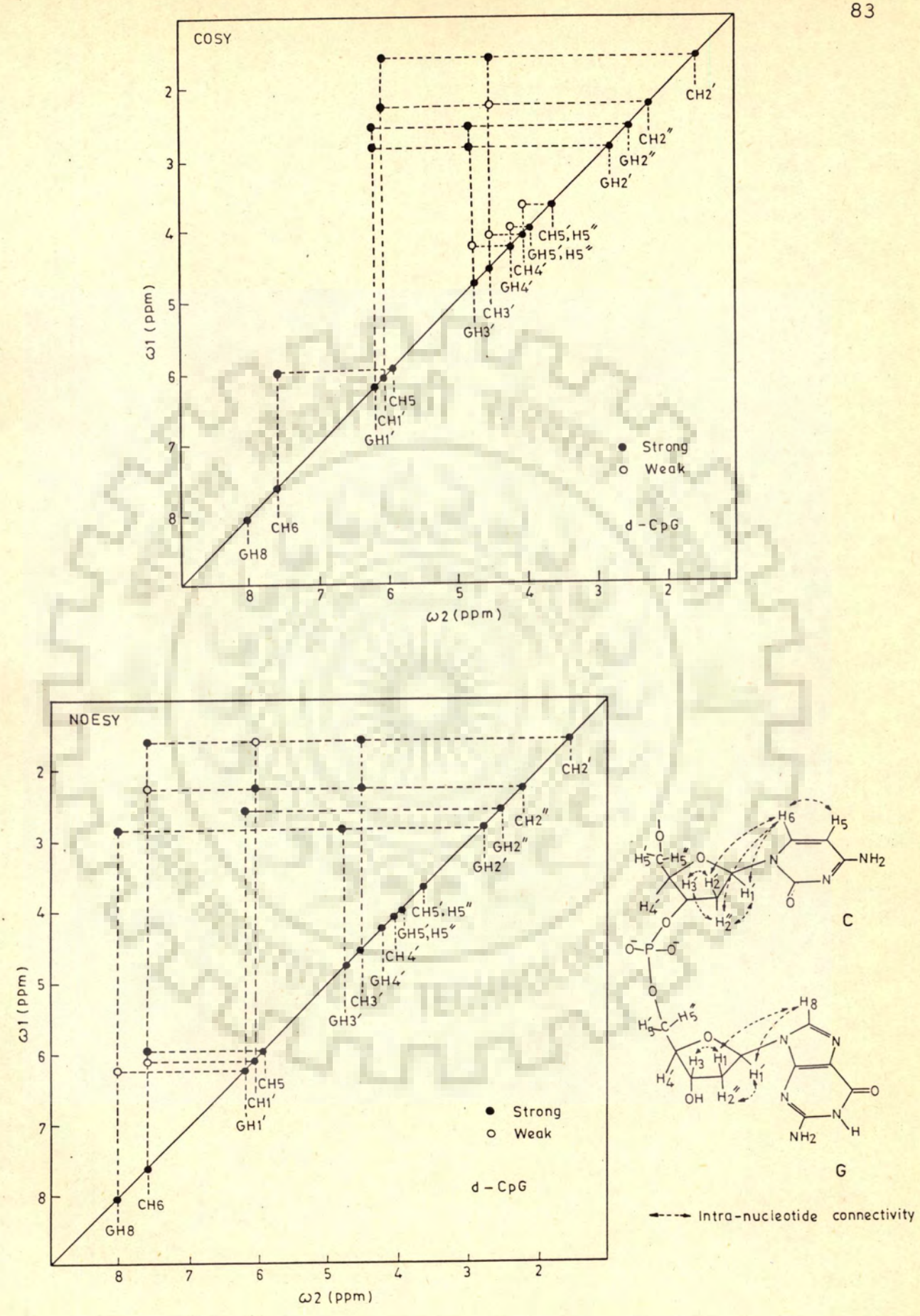


Fig. 3.6 Schematic representation of the results of COSY and NOESY spectra of d-CpG shown in Figs. 3.3 to 3.5.

nucleotide sugars exist in 01'-endo conformation. (Ref. Chap. 2 page 57).

The results on relative intensities of NOESY spectrum (Figure 3.6) are as follows:

GH8 ... GH2'	strong	(9)
GH8 ... GH2"	nil	(10)
GH8 ... GH1'	weak	(11)
GH1' ... GH2'	Nil	(12)
GH1' ... GH2"	strong	(13)
GH2' ... GH3'	strong	(14)
GH2" ... GH3'	nil	(15)
CH6 ... CH2'	strong	(16)
CH6 ... CH2"	weak	(17)
CH6 ... CH1'	weak	(18)
CH1'... CH2'	weak	(19)
CH1... CH2"	strong	(20)
CH2'...CH3'	strong	(21)
CH2"...CH3'	strong	(22)

The presence of strong intra-residue connectivity of base protons with corresponding H2' protons [(9) and (16)] and absence of intra-residue connectivity CH6...CH5" and inter-residue connectivities GH8...CH5", CH5' clearly indicate

the existence of right-handed DNA structure of d-CpG in solution. This is consistent with the findings of Viswamitra et al [31] who by X-ray diffraction studies found that d-CpG exists as right—handed double—helix with C2'— endo sugar conformation for both C and G bases and glycosidic torsion angles 55° for cytidine and 94° for guanosine. However, the same workers found that d-CpG when complexed with mitoxantrone adopts left-handed double-helix structure [unpublished].

The glycosidic bond rotation χ_{CN} defined as dihedral angle O1' - C1' - N1 - C6 in pyrimidines and O1' - C1' - N9 - C8 in purines is ascertained from NOE results. Since base to H2' and H2" connectivities are of different intensities and further base to H2" connectivity is weaker than that from base to H2' for both nucleotides, it is inferred that both G and C exist in anti glycosidic bond conformation.

We have obtained the interproton distances from the Dreiding model of d-CpG (Figure 3.7) in a standard B—DNA geometry of Arnott and Hukins [7] with glycosidic angle in anti conformation being equal to -90°/262° and torsional angles as ω = -46°, Φ = -147°, Ψ = 36°, Ψ' = 157°, Φ = 155°, ω' = -96° for both C2' endo as well as O1' endo sugar conformations [7]. The interproton distances of furanose ring thus obtained are given in Table 3.2. For comparison the distances obtained by Mazeau [90] and Gopalkrishnan, Bansal [49] are also shown in table. The only significant difference

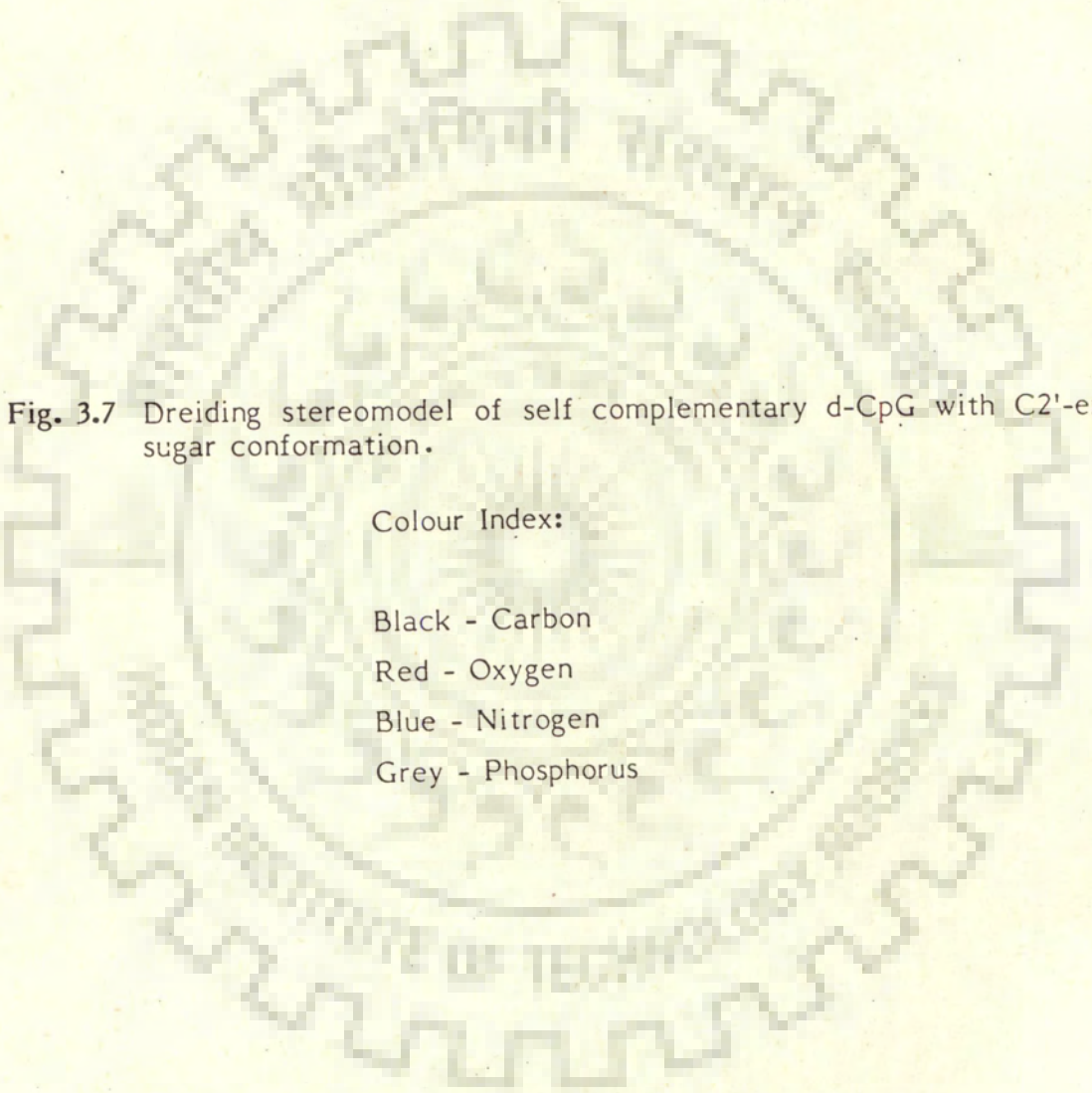


Fig. 3.7 Dreiding stereomodel of self complementary d-CpG with C2'-endo sugar conformation.

Colour Index:

Black - Carbon

Red - Oxygen

Blue - Nitrogen

Grey - Phosphorus

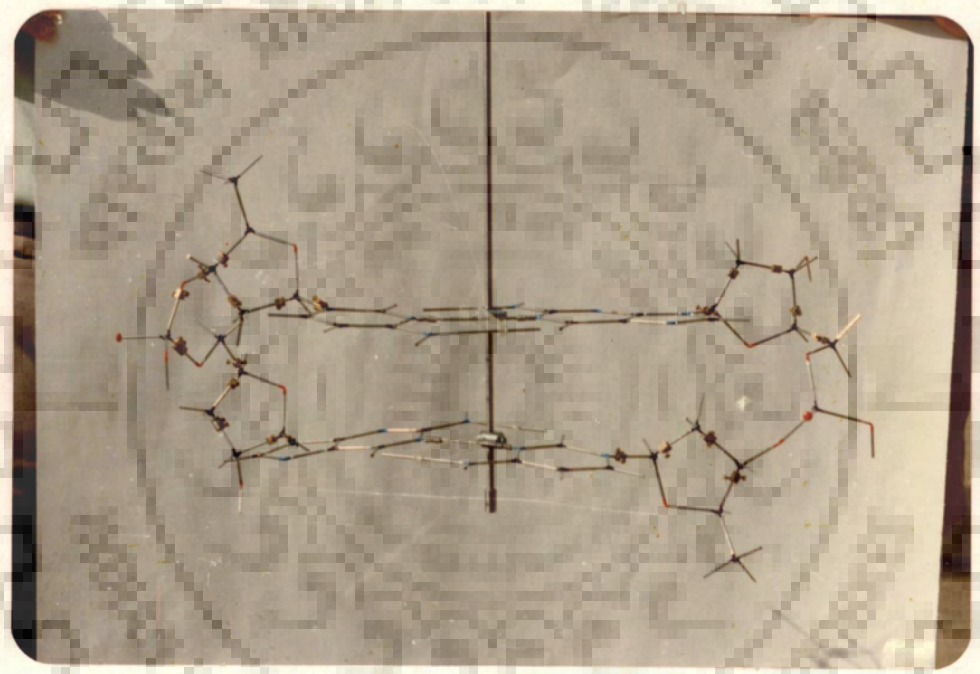


Table 3.2 : Comparison of interproton distances in d-CpG obtained from Dreiding stereo-models with that from other methods.
A. Furanose ring protons

Sl. No.	Inter-proton vector	O1'-/O4'-endo		C2'-endo		Gopal Krishnan and Bansal* [49]	2D NMR (COSY)
		Mazeau (in Å)	Model (in Å)	Mazeau (in Å)	Model (in Å)		
1	2	3	4	5	6	7	8
1.	H1'-H2"	2.19	G 2.2	2.26	G 2.4	2.3(2.6)	+
			C 2.3		C 2.3		
2.	H1'-H2'	2.81	G 2.9	2.62	G 2.5	3.0(3.1)	-
			C 2.9		C 2.6		
3.	H1'-H4'	3.8	G 3.5	3.83	G 3.8	3.7	-
			C 3.6		C 3.8		
4.	H1'-H3'	-	G 3.9		G 4.2		-
			C 4.0		C 4.2		
5.	H1'-H5'		G 4.8		G 4.6		-
			C 4.8		C 4.6		
6.	H1'-H5"		G 4.6		G 5.0		-
			C 4.7		C 5.0		
7.	H2'-H2"		G 1.8		G 1.8		+
			C 1.8		C 1.8		
8.	H2'-H3'		G 2.3		G 2.3	2.4	+
			C 2.4		C 2.3		
9.	H2'-H4'		G 4.2		G 4.3		-
			C 4.1		C 3.9		
10.	H2'-H5'		G 4.3		G 4.2		-
			C 4.3		C 4.3		
11.	H2'-H5"		G 5.0		G 5.0		-
			C 5.0		C 4.9		
12.	H2"-H3'		G 3.8		G 2.9	2.7	-
			C 3.9		C 3.0		
13.	H2"-H4'		G 3.4		G 3.9	3.9	-
			C 3.2		C 3.8		

Contd.

1	2	3	4	5	6	7	8
14.	H2"-H5'		G 4.8		G 4.7		-
			C 4.7		C 4.8		
15.	H2"- H5"		G 5.4		G 5.4		-
			C 5.4		C 5.6		
16.	H3'-H4'		G 2.1		G 2.8	2.7	-
			C 2.2		C 2.8		
17.	H3'-H5'		G 2.3		G 2.2		-
			C 2.3		C 2.2		
18.	H3'-H5"		G 3.5		G 3.6	2.8	-
			C 3.4		C 3.7		
19.	H4'-H5'		G 2.7		G 2.7		-
			C 2.7		C 2.7		
20.	H4'-H5"		G 2.4		G 2.4		-
			C 2.4		C 2.4		
21.	H5'-H5"		G 1.8		G 1.8		-
			C 1.8		C 1.8		

Note : The values within parentheses are listed by Reid et al [Reid, D.G., Salisbury, S.A. Bellard, S. Shakked,Z. and Williams, D.H. (1983) Biochemistry 22, 2019. for the refined B-Form model proposed by Arnott and Chandrasekharan.

The + and - sign indicate the presence and absence of cross peaks in the 2D-COSY spectra respectively.

between two sugar conformations is in interproton distances H3' - H4' which is less in O1'—endo conformation than that in C2'—endo conformation. It may be noted that a general agreement between the results obtained by us and that of other workers exists. The existence of NOE connectivities H1' - H2", H2' - H2", H2' - H3', H3' - H4', H4' - H5" and absence of connectivity H1' - H2' in the experimental NMR data on d-CpG, also indicated in table, are consistent with the results obtained by modelling studies.

The interproton distances between base and furnose protons obtained from Dreiding stereomodels (Figure 3.7) are given in Table 3.3. The distances obtained by Govil's method [52] using contour plots of distances as a function of pseudorotation and glycosidic bond rotation P = 90° and 162° for O1' endo and C2' endo conformations, respectively and χ_{CN} = 82° and that by Gopalkrishan,Bansal [49] are also tabulated alongwith observed NOE's. By comparing these results it is easily seen that experimental results agree with those obtained by modelling and other workers. Base to H2' intra-residue NOE's are seen in d-CpG and inter-residue NOE's

$$(\text{base proton})_i \ldots (\text{H2"})_{i-1}$$

and $$(\text{base proton})_i \ldots (\text{H1'})_{i-1}$$

are weaker than the intra-residue ones.

Table 3.3 : Comparison of interproton distances in d-CpG obtained from Dreiding stereo models with that from other methods. B. Base to furanose ring protons

Sl. No.	Inter-proton vector	O1'-/O4'-endo Govil (in Å)	O1'-/O4'-endo Model (in Å)	C2'-endo Govil (in Å)	C2'-endo Model (in Å)	Gopala Krishnan and Bansal* [49]	2D NMR (NOESY)
1	2	3	4	5	6	7	8
1	Base-H2'	2	G 1.9	2	G 1.9	2.2(2.3)	+
			C 1.6		C 1.7	2.4(2.0)	
2	Base-H2"	>3	G 3.5	>3	G 3.9	3.7(3.5)	-
			C 3.3		C 3.8	3.9 (3.3)	
3	Base-H3'	3.5	G 3.5	4	G 3.9	4.2	-
			C 3.0		C 3.7		
4	Base-H4'	4.5	G 5.1	4.5	G 5.1		-
			C 4.6		C 4.8		
5	Base-H1'		G 3.8		G 3.8	3.8(3.8)	-
			C 3.6		C.3.7	3.6(3.7)	
6	Base-Previous H1'		GH8-CH1' =4.1		GH8'-CH1' =3.9	2.8(3.6)	-
7	Base-Previous H2'		GH8-CH2' =3.2		GH8-CH2' =3.2	3.6(3.9)	-
8	Base-Previous H2"		GH8-CH2" =2.9		GH8-CH2" =3.0	2.2(2.1)	-
9	Base-Previous H3'		GH8-CH3' =4.8		GH8-CH3' =4.9	4.9	-
10	Base to Base		GH8-CH6 =5.0			5.2	-

Note : The values within parentheses are listed by Reid et al [Reid, D.G., Salisbury, S.A. Bellard, S. Shakked,Z. and Williams, D.H. (1983) Biochemistry, 22, 2019. for the refined B-Form model proposed by Arnott and Chandrasekharan.

The + and - sign indicate the presence and absence of cross peaks in the 2D-NOESY spectra.respectively.

Hence d-CpG adopts a B—DNA like conformation with O1'—endo sugar pucker and anti glycosidic bond rotation for both C as well as G residues.

BINDING OF d-CpG TO Lys-Trp-Gly-Lys OtBu

Figure 3.8 shows the results of binding of 5.0 mM (single strand concentration) d-CpG to tetrapeptide. 2.5 mM tetrapeptide has been taken at 500 MHz at three temperatures 288 K, 297 K and 355 K and it is found that the chemical shift of peptide protons does not change with temperature. The room temperature (297 K) spectra is shown in Figure 3.8 for reference. The resonance lines of various protons of Gly, Trp, and Lys have been assigned by selective decoupling experiments. Figure 3.9 shows a 500 MHz 2D—COSY spectra of 2.5 mM Lys-Trp-Gly-Lys OtBu at 297 K. The cross peaks indicating J-correlated connectivities and the chemical shift values of amino acid residues in peptides available in literature [146, 147] completely and unambiguously assign every NMR peak. However, the two lysines, that is, in the beginning and terminal end of tetrapeptide are indistinguishable from these results and have been named 1 and 4 arbitrarily. Figure 3.10 shows the chemical shifts of base and sugar protons of d-CpG and peptide protons on binding of d-CpG to Lys-Trp-Gly-Lys OtBu as a function of temperature. The values are taken from NMR spectra shows in Figure 3.8 and are also tabulated in Table 3.4. It is clearly seen

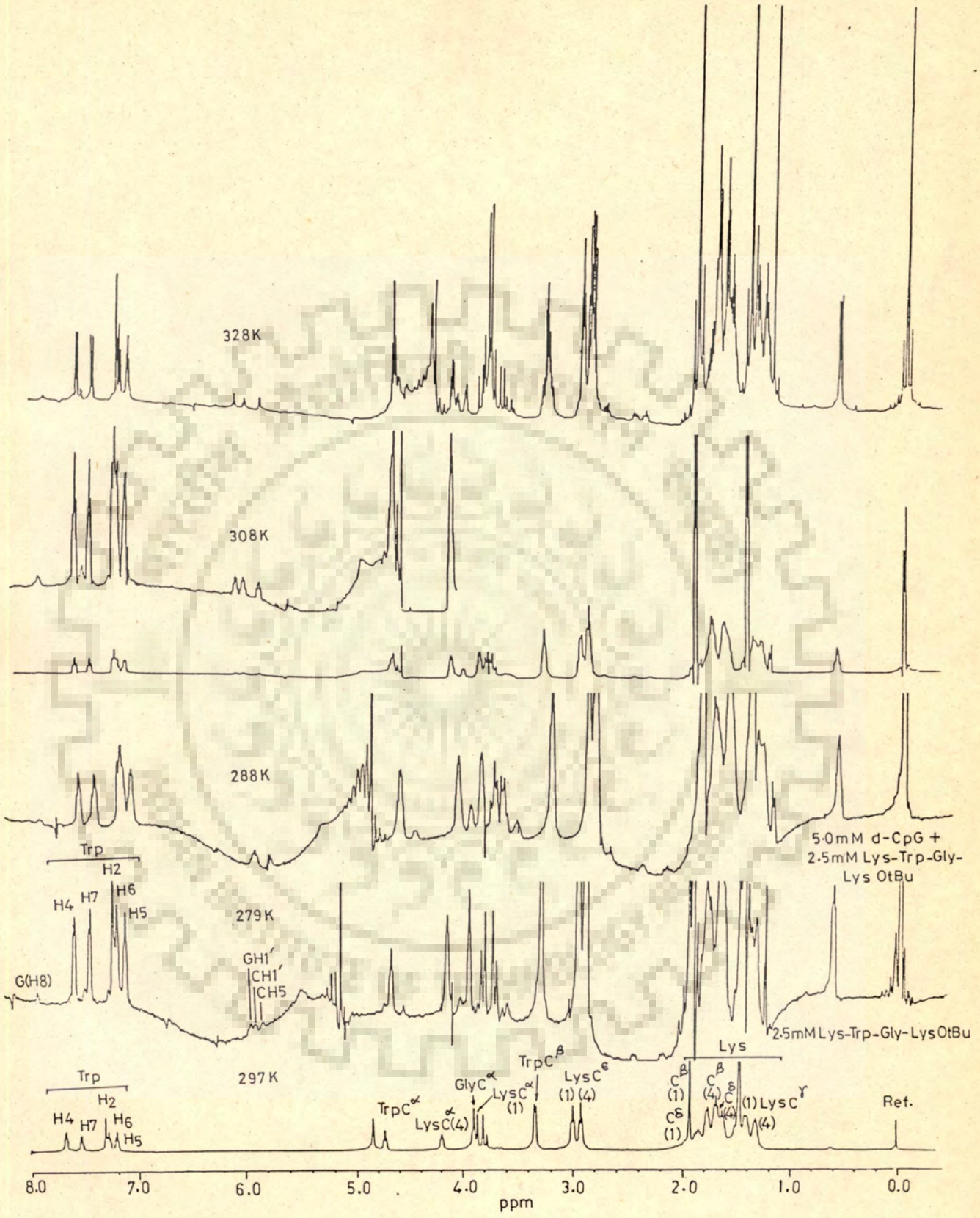


Fig. 3.8 The 500 MHz proton NMR spectra of a mixture of 5.0 mM d-CpG and 2.5 mM Lys-Trp-Gly-Lys OtBu at different temperatures in D_2O (pH=7.1) containing 0.25 mM EDTA. For comparison 500 MHz NMR spectrum of 2.5 mM Lys-Trp-Gly-Lys OtBu in D_2O solution at 297 K is also shown. Ref. DSS.

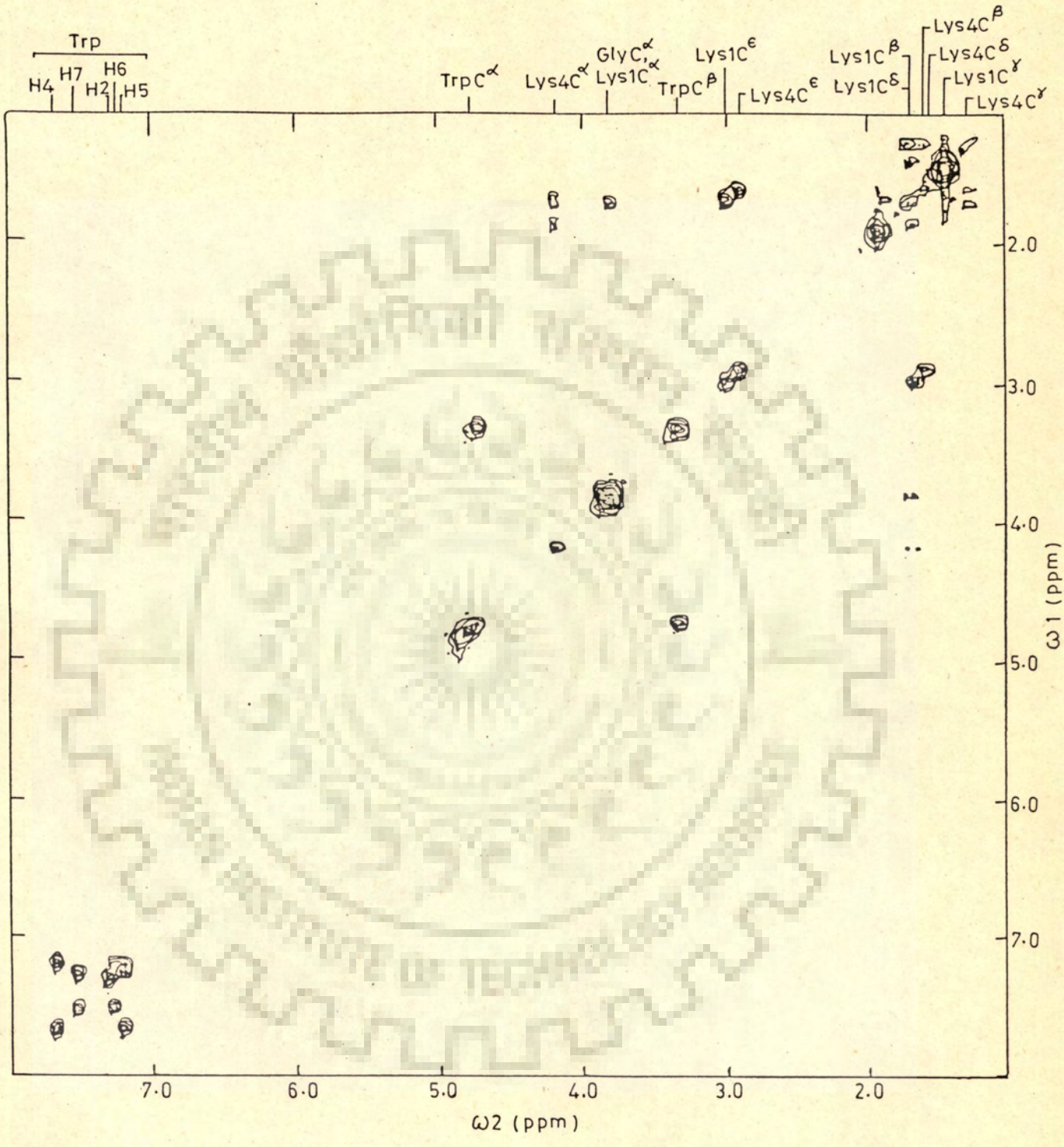


Fig. 3.9 500 MHz COSY spectrum of 2.5 mM Lys-Trp-Gly-Lys OtBu solution in D_2O at 297 K Ref. DSS.

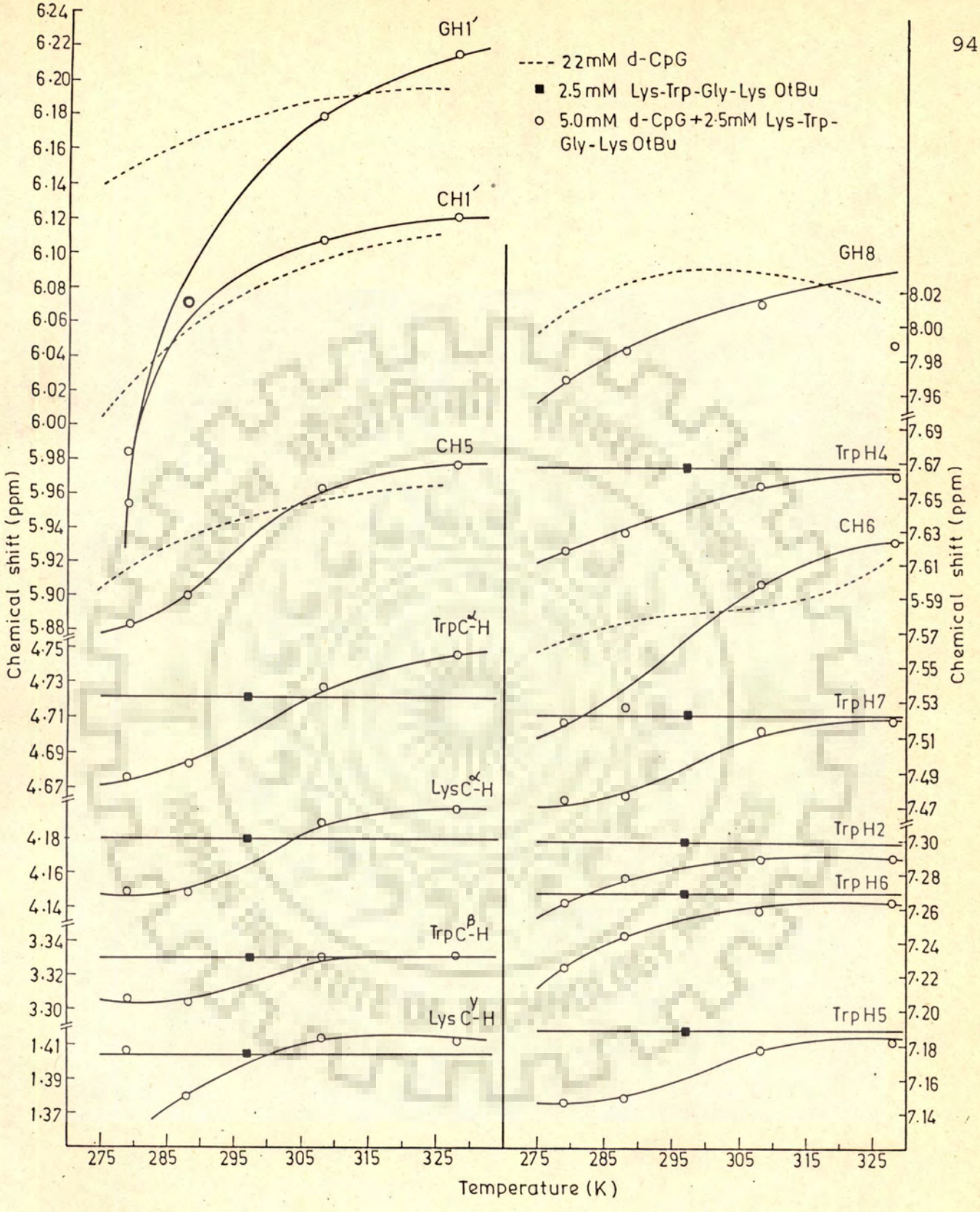


Fig. 3.10 Chemical shifts of some of the nucleotide and peptide protons as a function of temperature in uncomplexed nucleotide/peptide and in complexed state taken from data in Fig. 3.8.

Table 3.4 : Changes in chemical shift, $\Delta\delta$ in ppm, of peptide and nucleotide protons on binding of 5 mM d-CpG to 2.5 mM Lys-Trp-Gly-Lys OtBu at indicated temperatures

SYSTEM	Temp. (K)	GH8	CH6	CH5	GH1'	CH1'
	277	8.008	7.562	5.903	6.185	6.015
d-CpG	297	8.032	7.578	5.944	6.216	6.080
	310	8.027	7.589	5.957	6.229	6.101
	277	7.962	7.519	5.882	-	-
d-CpG + Lys-Trp-Gly-Lys	297	8.020	7.564	5.945	6.215	6.090
	310	8.018	7.598	5.964	6.184	6.108
	277	0.046	0.043	0.021	-	-
$\Delta\delta$(ppm)*	297	0.012	0.014	-0.001	0.001	-0.010
	310	0.009	-0.009	-0.007	0.045	-0.007

SYSTEM	Temp. (K)	Trp H4	Trp H2	Trp H7	Trp H6	Trp H5	Trp C^{α}-H	Trp C^{β}-H	Lys C^{α}-H (1)	Lys C^{α}-H (4)	Gly C^{α}-H
Lys-Trp-Gly-Lys	297	7.668	7.300	7.524	7.270	7.190	9.721	3.330	3.822	4.181	3.882
	277	7.614	7.258	7.470	7.218	7.146	4.672	3.302	-	4.146	-
d-CpG + Lys-Trp-Gly-Lys	297	7.644	7.284	7.499	7.256	7.162	4.702	3.324	3.824	4.166	3.857
	310	7.656	7.289	7.513	7.259	7.178	4.726	3.330	-	4.190	-
	277	0.054	0.042	0.054	0.052	0.044	0.049	0.028	-	0.032	-
$\Delta\delta$(ppm)*	297	0.024	0.016	0.030	0.014	0.028	0.019	0.006	-0.002	0.015	0.025
	310	0.012	0.011	0.011	0.011	0.012	-0.005	0.000	-	-0.009	-

Contd.

SYSTEM	Temp. (K).	Lys C^{β}-H		Lys C^{γ}-H		Lys C^{δ}-H		Lys C^{ε}-H	
		(1)	(4)	(1)	(4)	(1)	(4)	(1)	(4)
Lys-Trp-Gly-Lys	297	1.755	1.680	1.405	1.310	1.467	1.623	2.986	2.911
d-CpG + Lys-Trp-Gly-Lys	277	1.761	1.681	1.407	1.335	-	1.627	2.979	2.899
	297	1.771	1.682	1.407	1.300	1.470	1.620	2.983	2.912
	310	1.793	1.793	1.414	1.288	-	2.617	2.986	2.920
$\Delta\delta$ (ppm)*	277	-0.006	-0.001	-0.002	-0.025	-	-0.004	0.007	0.012
	297	-0.016	-0.002	-0.002	0.010	-0.003	0.003	0.003	-0.001
	310	-0.038	-0.003	-0.009	0.022	-	0.006	0.000	-0.009

*Upfield shifts, $\Delta\delta$ in ppm, are taken with positive sign.

that nucleotide base protons exhibit much more pronounced melting transition than that in the absence of tetrapeptide though the $T_{\frac{1}{2}}$ value is about the same that is, 295 - 297 K. The increase in difference of chemical shift at 330 K and at 277 K of proton resonances of d-CpG on addition of Lys-Trp-Gly-Lys OtBu is an indication that the double-helix has got more stabilised in presence of oligopeptide. The most significant change in chemical shift on binding occurs in ring protons of both, tetrapeptide and nucleotide. Practically all ring protons have shifted upfield. The GH8, CH6 and CH5 protons shift upfield by about 0.02 - 0.04 ppm at 277 K. The Tryptophan ring protons H2, H5, H4, H6 and H7 shift upfield by as much as ~0.05 ppm at 277 K. Further the change in chemical shifts, $\Delta\delta$, on binding decreases with temperature, it being large at 277 K than that at 320 K. This indicates that binding of tetrapeptide is more at lower temperature and hence greater to the double helical d-CpG than to the single-strand and hence is preferential. The upfield shift in Trp ring protons is attributed to stacking of Tryptophan ring with C-G base pairs of d-CpG. The same behaviour has been earlier reported on stacking or intercalation of aromatic rings of peptides or drugs in deoxynucleotide base pairs [42, 58, 33, 89, 108]. The change in chemical shift of lysine side chain protons particularly, the larger change in Lys C^{ε} protons, is likely due to the electrostatic interaction with negatively charges on the phosphates of DNA on binding of tetrapeptide to d-CpG.

Figures 3.11 and 3.12 are the results of 500 MHz COSY and NOESY experiments on the mixture of 5.0 mM d-CpG and 2.5 mM Lys-Trp-Gly-Lys OtBu at 297 K for the same sample solution and under identical solution conditions as in Figure 3.8. Assignments of various NMR peaks to specific protons of nucleotide and tetrapeptide are easily made by using Figures 3.3, 3.4 and 3.9 as a reference uncomplexed spectra. The results on relative intensities of cross peaks in COSY spectra, schematically in Figure 3.13, are as follows:

GH1' ... GH2' and GH2"	both equally strong	(1)
GH2' ... GH3'	strong	(2)
GH2" ... GH3'	nil	(3)
GH3' ... GH4'	weak	(4)
CH1' ... CH2' and CH2"	both equally strong	(5)
CH2' ... CH3'	strong	(6)
CH2" ... CH3'	nil	(7)
CH3' ... CH4'	nil	(8)

Using strategies discussed earlier it is clear that G sugar is in O1'-endo conformation whereas C sugar is in C2'-endo conformation in the compléxed d-CpG. The relative intensities in NOESY spectra (Figure 3.13) are:

GH8 ... GH2'	strong	(9)
GH8 ... GH2"	nil	(10)

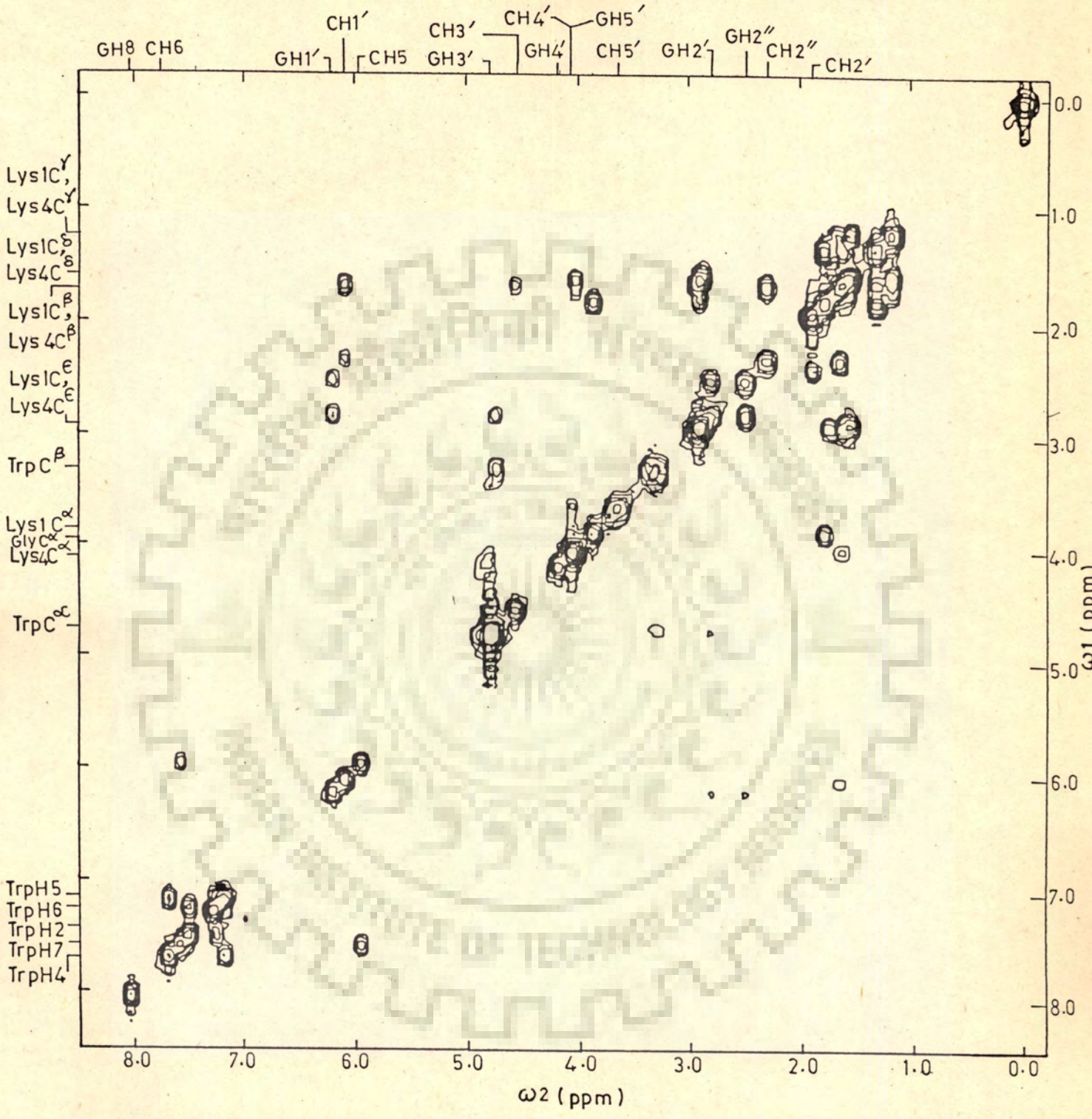


Fig. 3.11 500 MHz COSY spectrum of a mixture of 5.0 mM d-CpG and 2.5 mM Lys-Trp-Gly-Lys OtBu in D_2O, at 297 K for the sample used in Fig. 3.8.

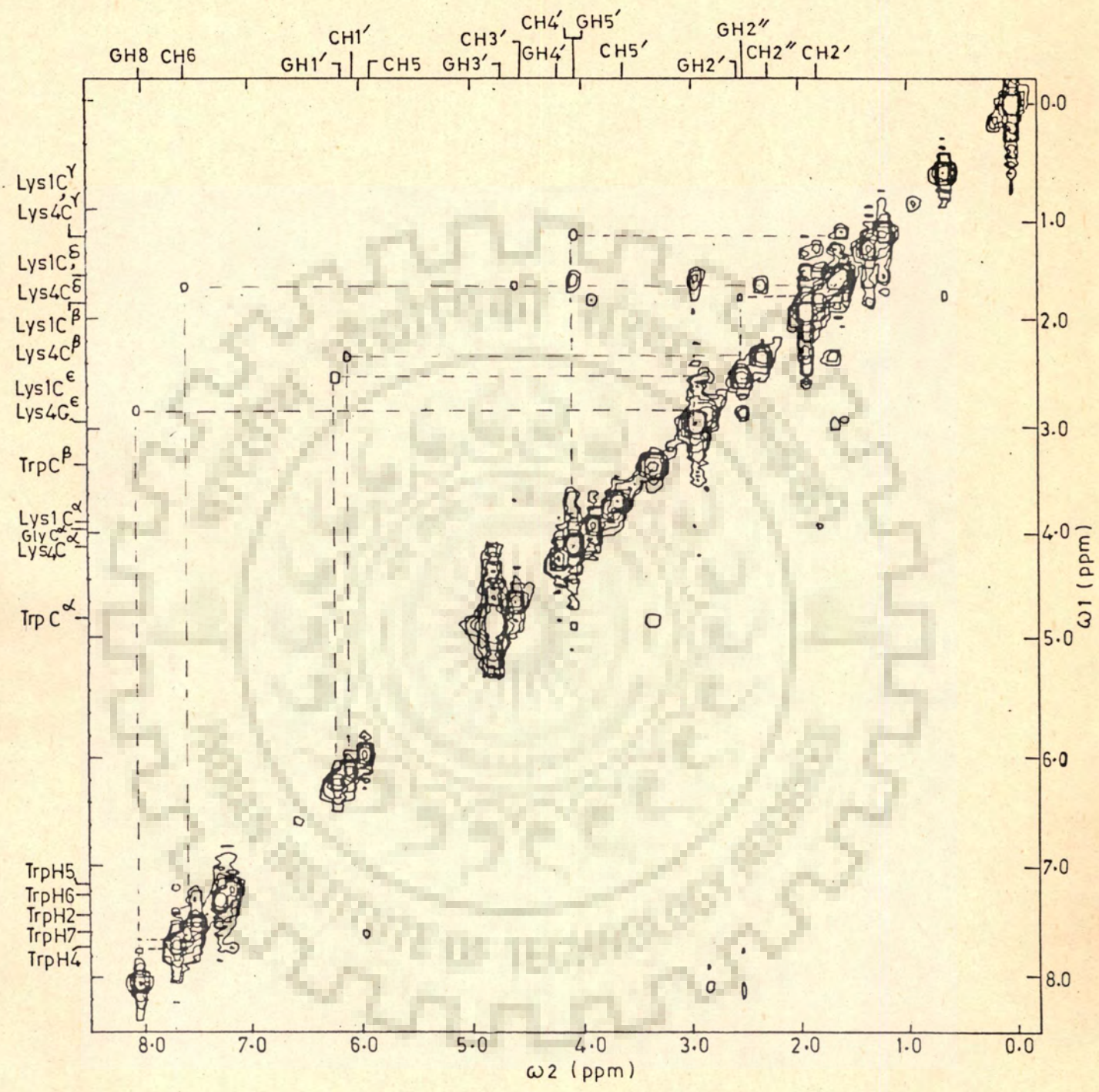


Fig. 3.12 500 MHz NOESY spectrum of a mixture of 5.0 mM d-CpG and 2.5mM Lys-Trp-Gly-Lys OtBu at 297 K for the same sample used in fig.3.8.

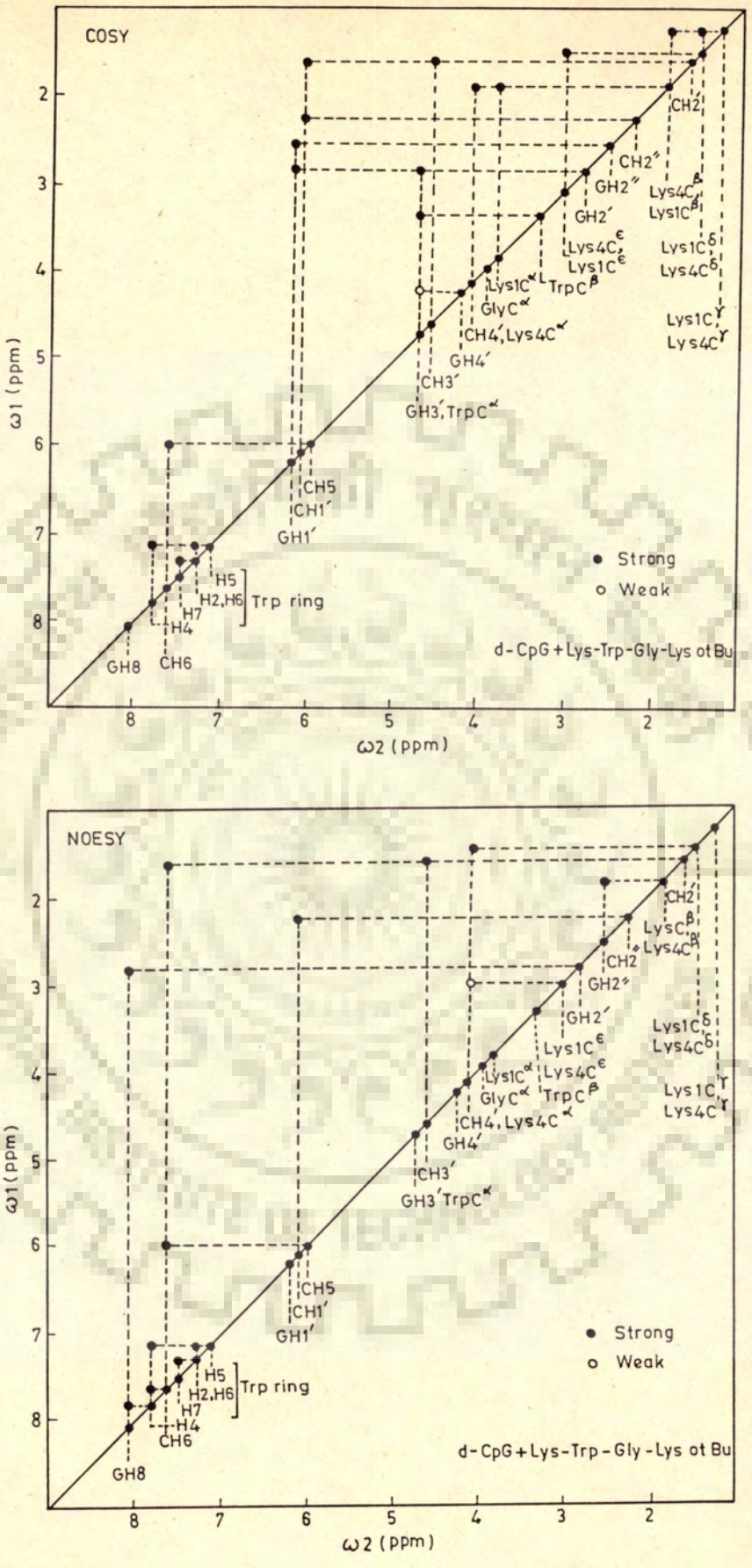


Fig. 3.13 Schematic representation of the results of COSY and NOESY spectra of a mixture of d-CpG with Lys-Trp-Gly-Lys OtBu shown in Figs. 3.11 and 3.12.

GH1' ... GH2'	nil	(11)
GH1' ... GH2"	strong	(12)
GH2' ... GH3'	nil	(13)
GH2" ... GH3'	nil	(14)
CH6 ... CH2'	strong	(15)
CH6 ... CH2"	nil	(16)
CH1' ... CH2'	nil	(17)
CH1' ... CH2"	strong	(18)
CH2' ... CH3'	strong	(19)
CH2" ... CH3'	nil	(20)

The presence of base proton CH8/CH6 intra-residue connectivity with H2' sugar proton and absence of any cross peak of base protons with H5" protons indicates the existence of right-handed DNA structure. Further, the stronger cross peak of base proton to H2' [(9) and (15)] compared to H2" (cross peak not seen) indicates anti glycosidic bond rotation for both residues.

Hence, the only change observed in conformation of d-CpG on binding to Lys-Trp-Gly-Lys OtBu is that the sugar conformation of residue C changes from O1'-endo to C2'-endo.

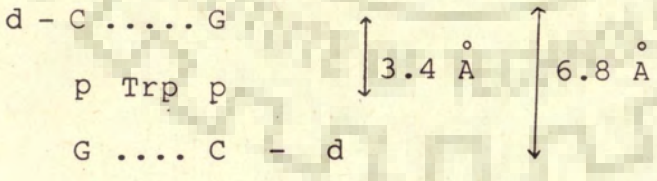
The NOESY spectrum of complex yields additional significant cross peaks between peptide protons and nucleotide

protons which define a specific geometry of complex. These connectivities are

NOE's-pairs of protons	Distance from Model	
GH8 ... Trp ring protons	3.4 Å	(21)
CH6 ... Trp ring protons	3.2 Å	(22)
GH2" ... Lys C^{β}		(23)
Lys C^{α}... Lys C^{γ} and Lys C^{ε}		(24)

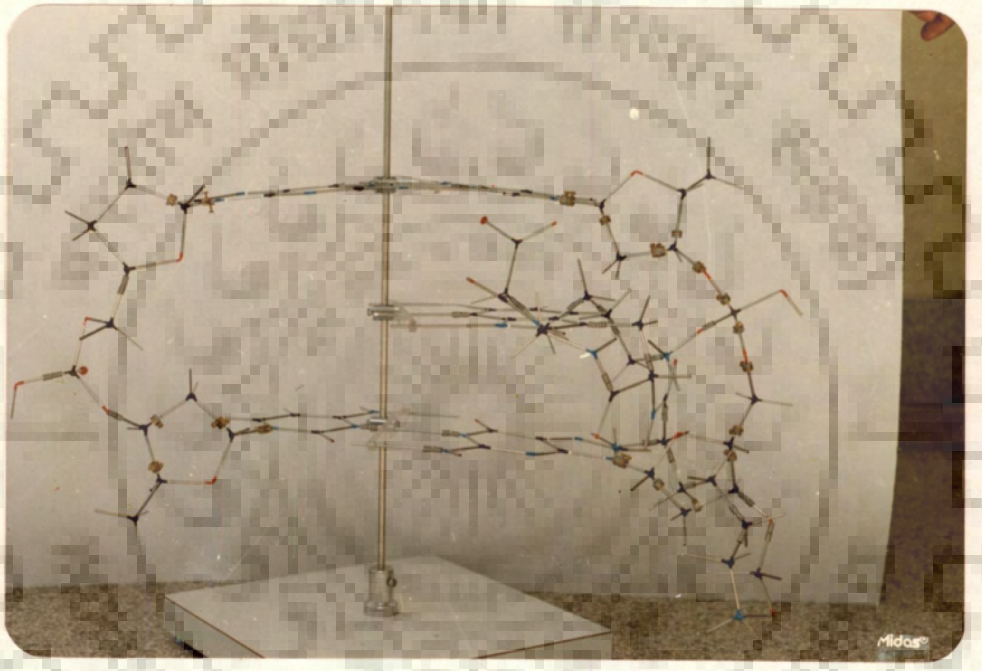
It is important to note that the existence of NOE No. (21) and (22) demonstrate that Trp ring intercalates between C-G base pairs.

A Dreiding stereomodel of complex of d-CpG double helix with tetrapeptide Lys-Trp-Gly-Lys OtBu has been constructed to intercalate Trp ring within the base pairs as shown below:



A view of the model is shown in Figure 3.14. On increasing the distance between base pairs from 3.4 to 6.8 Å, the backbone torsional angles have been changed from standard B-DNA [7] structure as follows:

Fig. 3.14 Dreiding stereomodel of complex of d-CpG with Lys-Trp-Gly-Lys OtBu showing intercalation of Trp between base pairs.



Torsional angle	χ	Φ'	ω'	ω	Φ	ψ
B-DNA	-98°	155°	-96°	-46°	-147°	36°
d-CpG+Lys-Trp-Gly-Lys. model	-98°	-115°	+210°	-190°	-140°	170°

The position of Trp ring in a plane parallel to base pairs has been so adjusted so as to result in change in chemical shifts as observed experimentally (that is, Table 3.4) and conforms with the magnetic anisotropy calculation of Pullman and Coworkers [47]. The distances between pairs of protons for which NOE's are experimentally seen in complex that is no. (21) to (24) above have been obtained from the model and are shown in parenthesis against NOE's no. (21) to (24). The results are in qualitative agreement and suggest that a specific geometry of complex may possibly exist.

BINDING OF d-CpG to Lys-Tyr-Lys

Figures 3.15 and 3.16 show the results of binding of 5.0 mM d-CpG to 2.5 mM tripeptide Lys-Tyr-Lys at different temperatures. The proton NMR spectra of tripeptide at 297K is also shown below in Figure 3.15 for reference. It has been verified that peptide proton chemical shifts do not show changes with temperature in the range 285 - 355 K.

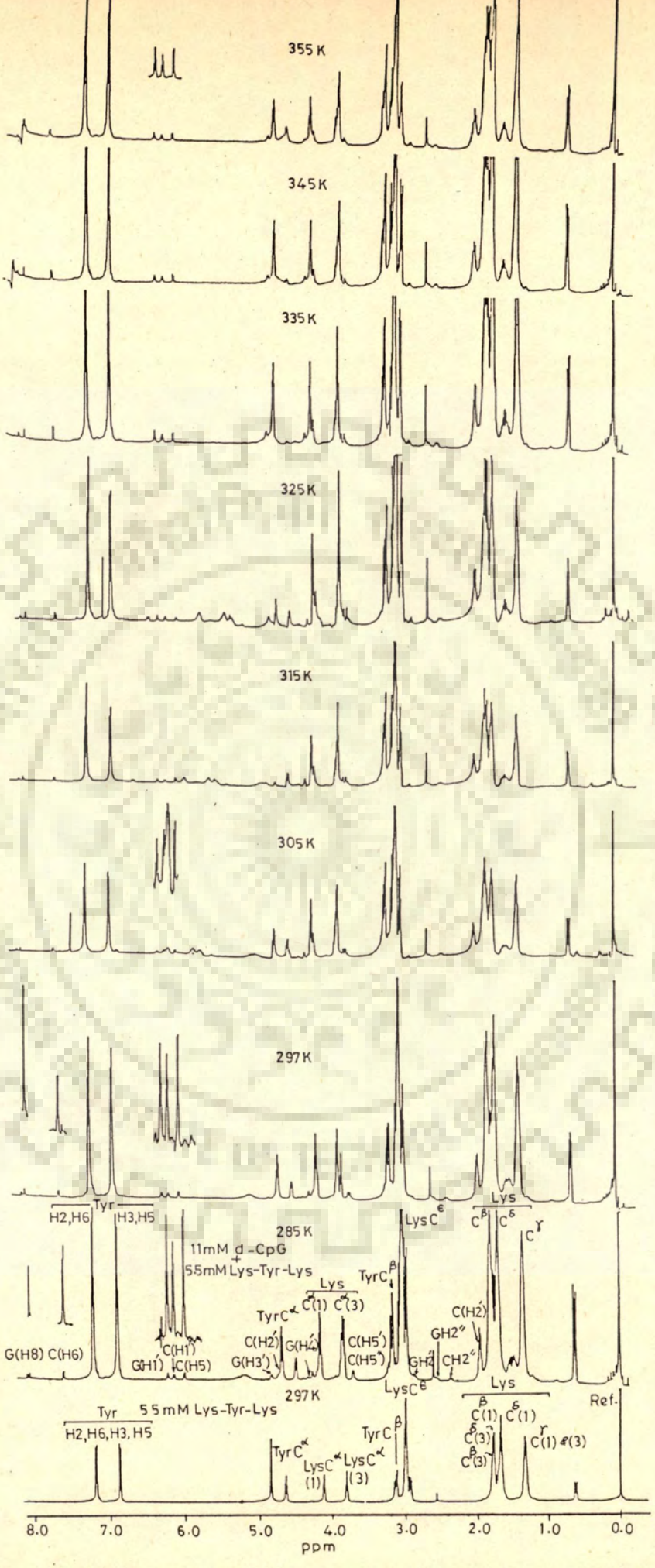


Fig. 3.15 500 MHz proton NMR spectra of a mixture of 11.0 mM d-CpG and 5.5 mM Lys-Tyr-Lys at different temperatures in D_2O (pH=7.1) containing 0.25 mM EDTA. 500 MHz NMR spectrum of 5.5mM Lys-Tyr-Lys

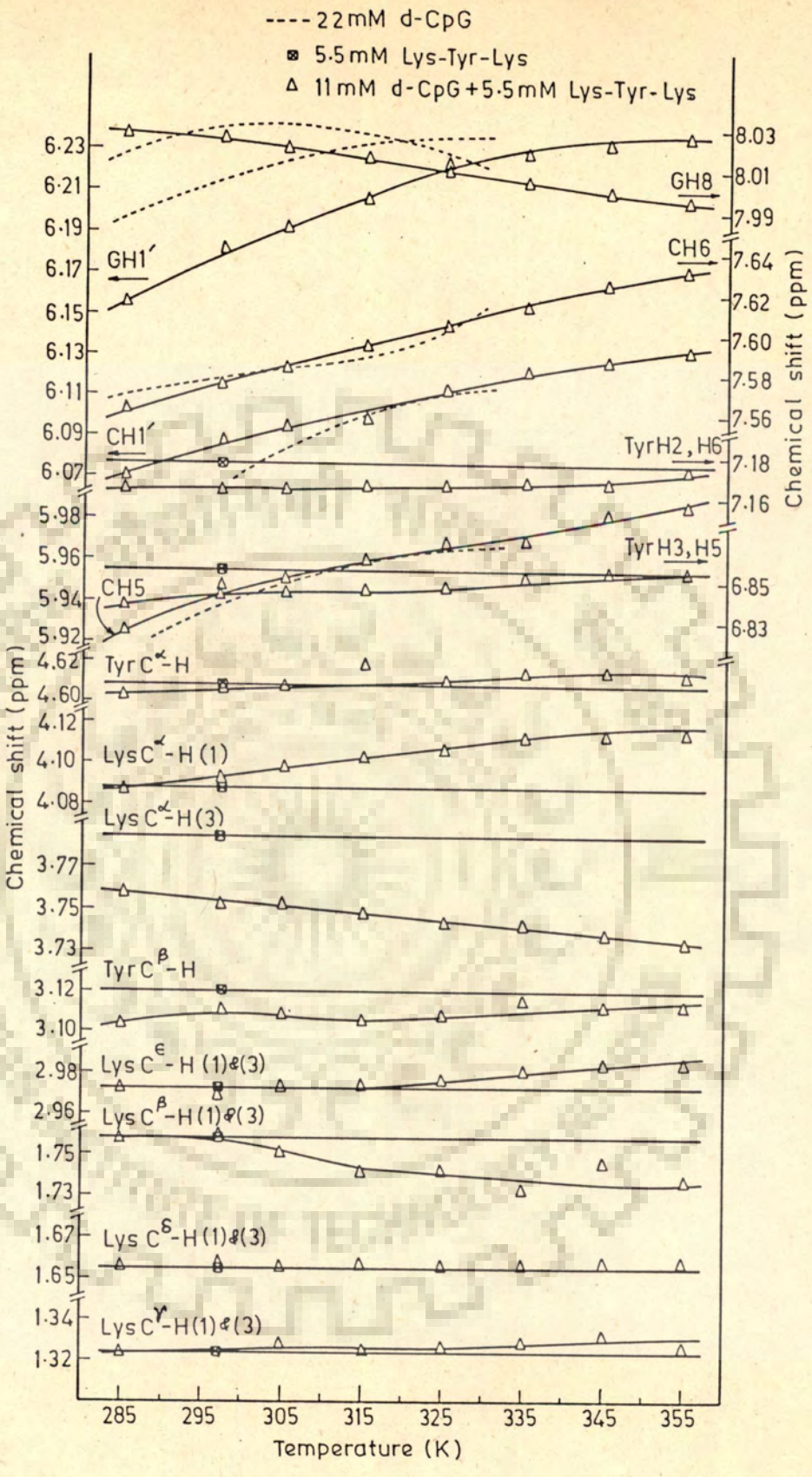


Fig. 3.16 Chemical shifts of some of the nucleotide and peptide protons as a function of temperature in the uncomplexed nucleotide/peptide and in complexed state taken from data in Fig. 3.15.

The specific assignments in peptide spectra have been made using selective decoupling experiments and the δ values of amino acid residues in peptides [146, 147]. The two lysines are indistinguishable. The chemical shift changes on binding are tabulated in Table 3.5. Changes in chemical shift for all protons are larger at lower temperatures, i.e. at 285 K than that at 355 K and decrease with temperature. It indicates that binding may preferentially occur to double-stranded deoxyoligonucleotide than the single-strand. $T_{\frac{1}{2}}$ value of d-CpG on binding increases from 295 K to 305-322 K. CH6 and CH5 protons shift downfield as though partial destacking has occurred whereas the Tyrosine ring protons shift upfield by ~0.02 ppm at 285 K. The Tyrosine ring may partially stack between base pairs and hence result in upfield shifts observed. However, it may be noted that the upfield shifts on binding to Lys-Tyr-Lys are lesser than the corresponding shifts on binding to Lys-Trp-Gly-Lys OtBu.

Figures 3.17 and 3.18 are the results of COSY and NOESY experiments at 297 K on mixture of d-CpG with Lys-Tyr-Lys for the same sample solution and under identical conditions as in Figure 3.15. The results on relative intensities of cross peaks in COSY spectra (See Figure 3.19) are:

GH1' ..... GH2' and GH2"	both equally strong	(1)
GH2' ..... GH3'	weak	(2)
GH2" ..... GH3'	nil	(3)

Table 3.5 : Changes in chemical shift, $\Delta\delta$ in ppm, of peptide and nucleotide protons on binding of 11 mM d-CpG to 5.5 mM Lys-Tyr-Lys at indicated temperatures

SYSTEM	Temp. (K)	GH8	CH6	CH5	GH1'	CH1'	Tyr H2,H6	Tyr H3,H5
d-CpG	285	6.025	7.571	5.926	6.202	6.052	-	-
Lys-Tyr-Lys	297	8.032	7.578	5.945	6.216	6.080	7.177	6.855
	315	8.025	7.593	5.960	6.233	6.105	-	-
d-CpG + Lys-Tyr-Lys	285	8.028	7.562	5.926	6.126	6.070	7.165	6.838
	297	8.025	7.575	5.943	6.181	6.088	7.164	6.657
	315	8.015	7.594	5.960	6.206	6.098	7.165	6.845
$\Delta\delta$(ppm)*	285	-0.003	0.009	-	0.046	-0.018	0.012	0.017
	297	0.007	0.016	0.002	0.03^5	-0.008	0.013	0.008
	315	0.01^0	-0.001	-	0.027	0.007	0.012	0.012

SYSTEM	Temp. (K)	$^{\alpha}$TyrC-H	$^{\beta}$TyrC-H	Lys (1)	$^{\alpha}$C-H (3)	$^{\beta}$LysC-H (1)&(3)	$^{\gamma}$LysC-H (1)&(3)	$^{\delta}$LysC-H (1)&(3)	$^{\varepsilon}$LysC-H (1)&(3)
Lys-Tyr-Lys	297	4.608	3.121	4.089	3.785	1.758	1.324	1.655	2.971
d-CpG + Lys-Tyr-Lys	285	4.603	3.104	4.087	3.758	1.759	1.323	1.656	2.971
	297	4.608	3.111	4.092	3.751	1.759	1.321	1.658	2.970
	315	4.618	3.105	4.103	3.748	1.741	1.324	1.656	2.975
$\Delta\delta$ (ppm)*	285	0.005	0.017	0.002	0.027	-0.001	0.001	-0.001	-
	297	-	0.010	-0.003	0.034	-0.001	0.003	-0.003	0.001
	315	-0.01^0	0.016	-0.014	0.037	0.017	0.000	-0.001	-0.004

*Upfield shifts, $\Delta\delta$ in ppm, are taken with positive sign.

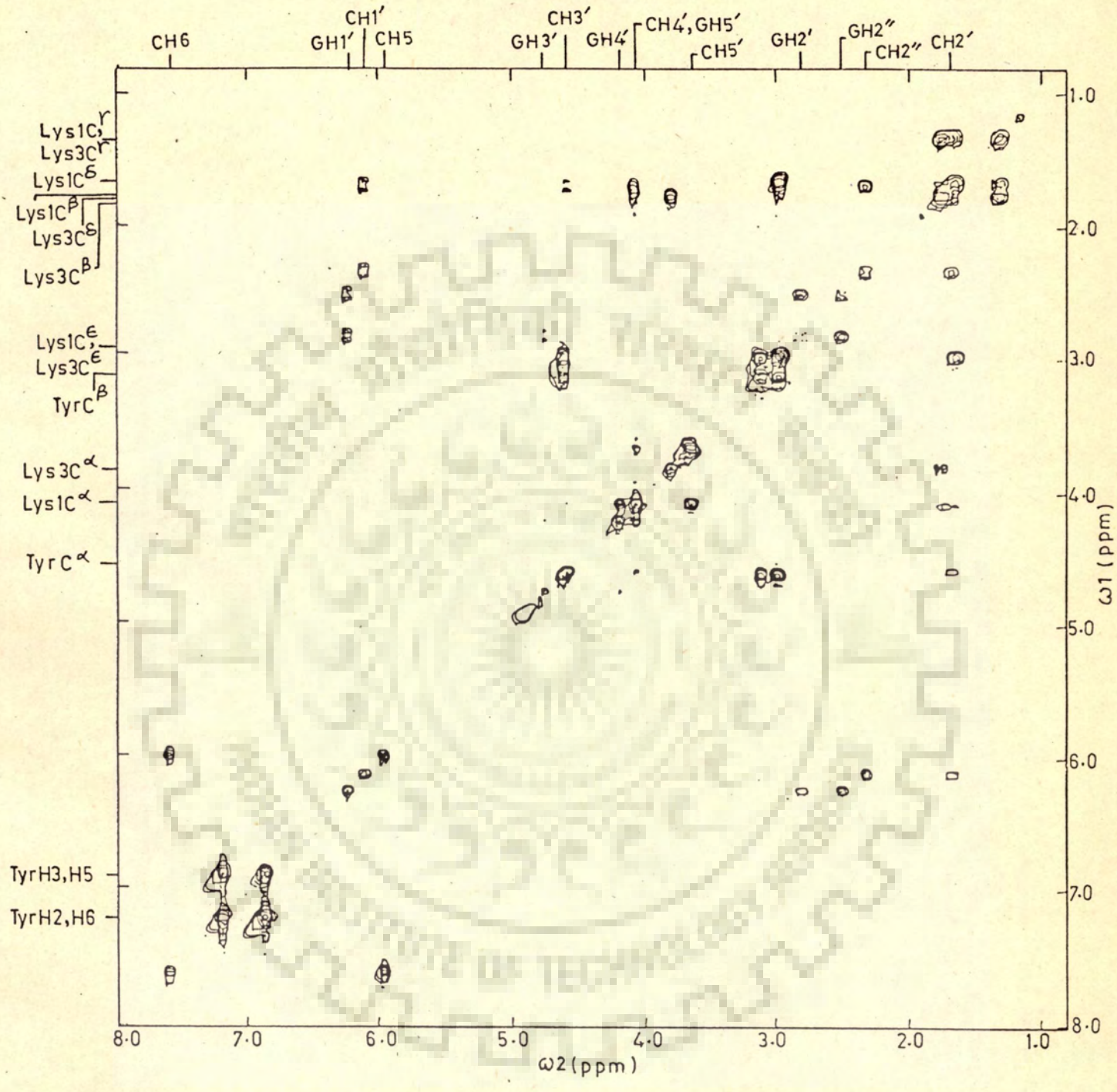


Fig. 3.17 500 MHz COSY spectrum of a mixture of 5.0 mM d-CpG and 0.125 mM Lys-Tyr-Lys in D_2O (pH=7.0) containing 0.25 mM EDTA at 297 K.

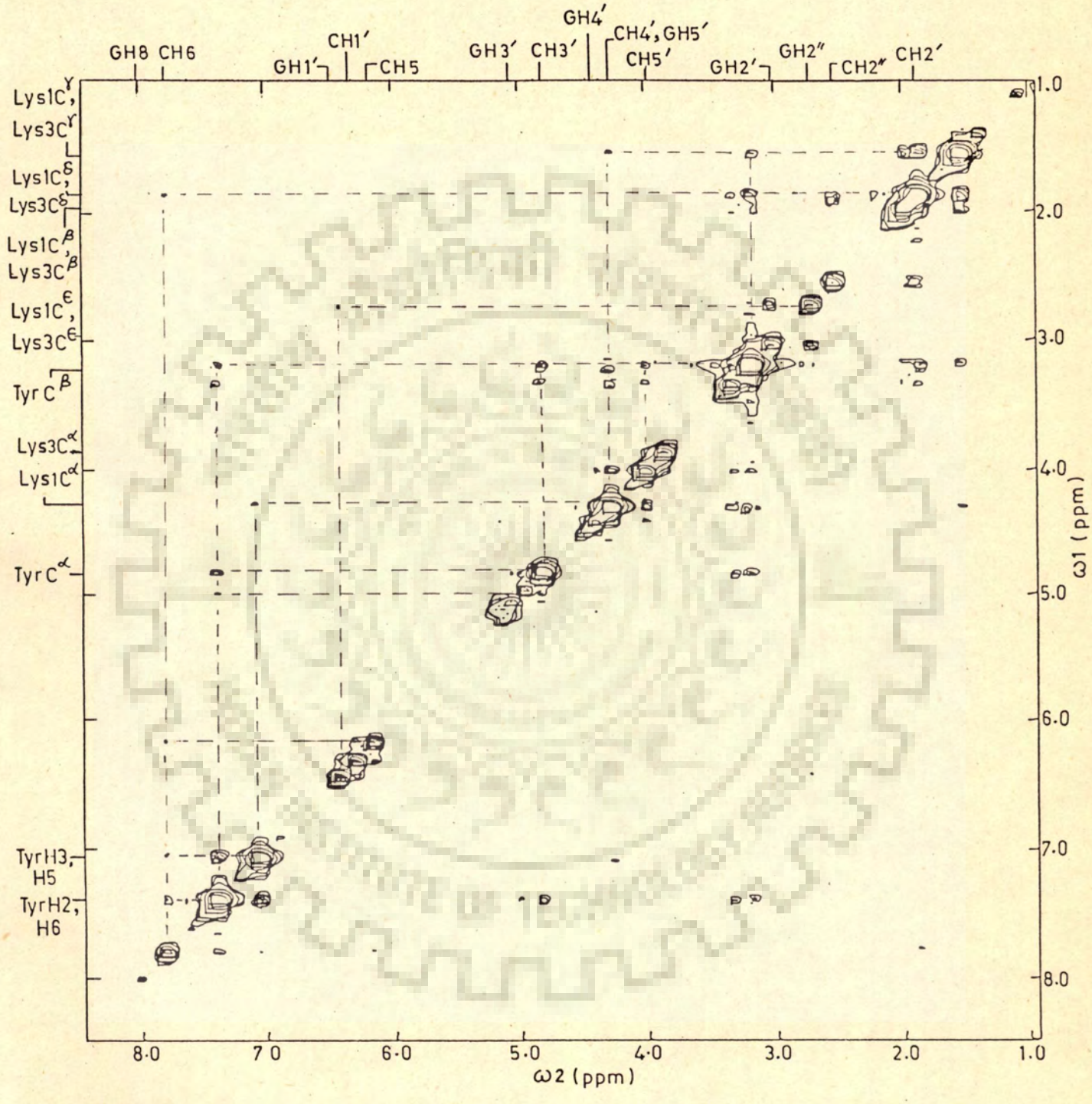
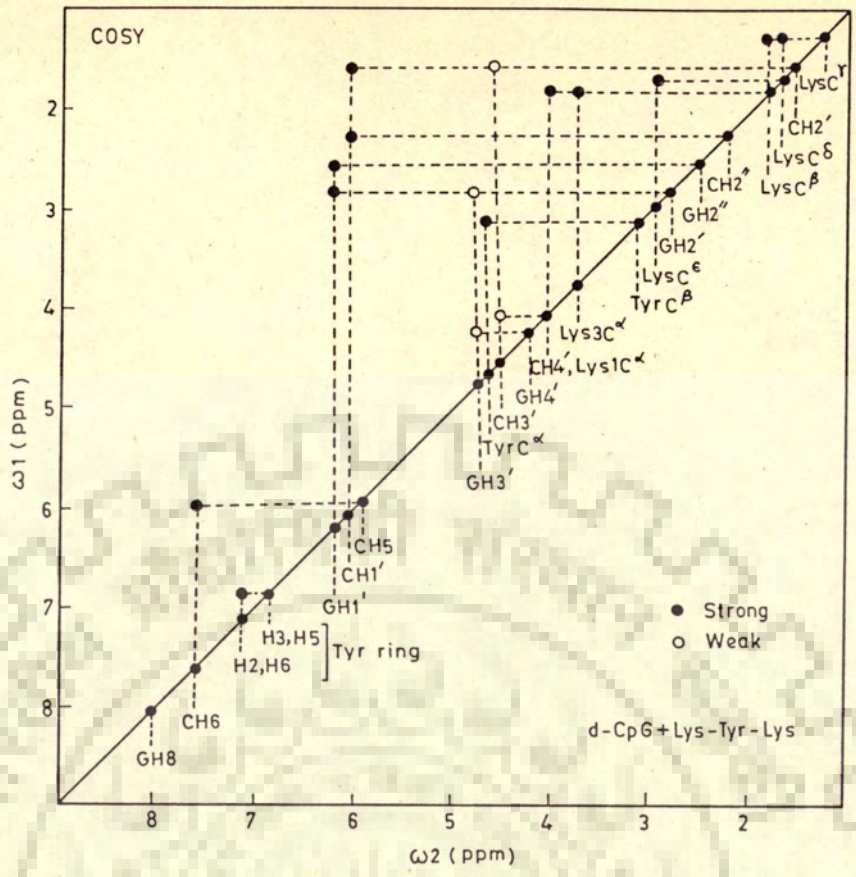


Fig. 3.18 500 MHz NOESY spectrum of a mixture of 5.0 mM d-CpG and 2.5 mM Lys-Tyr-Lys in D_2O (pH=7.0) containing 0.25 mM EDTA at 297 K.



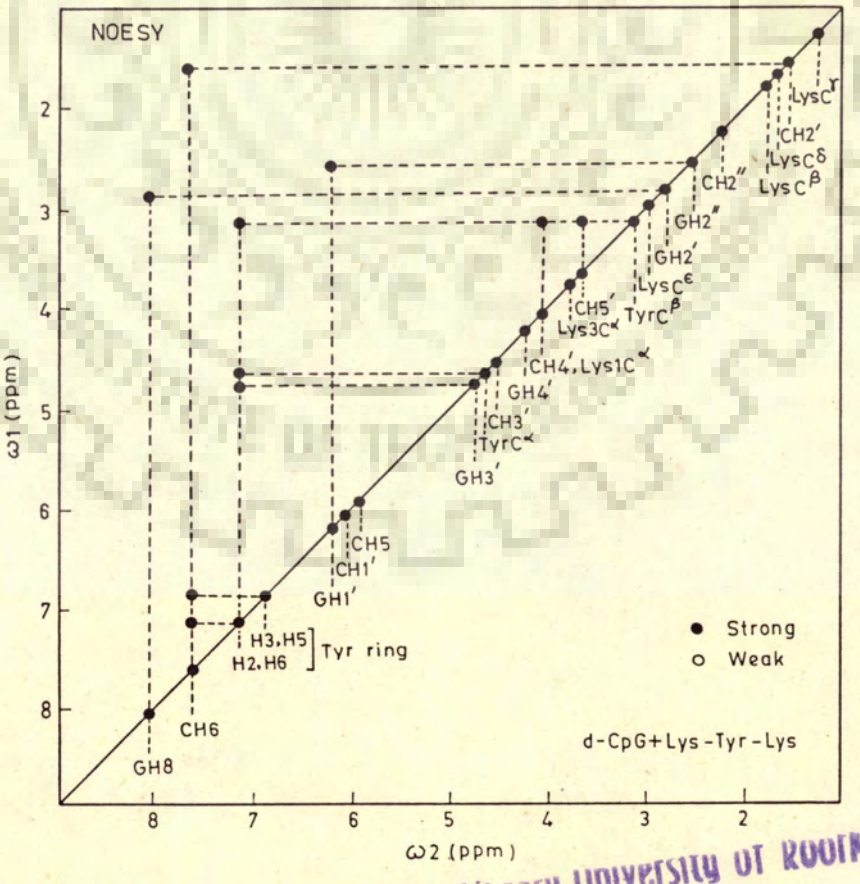


Fig. 3.19 Schematic representation of the results of COSY and NOESY spectra of a mixture of d-CpG with Lys-Tyr-Lys shown in Fig.3.17-18

GH3"	GH4'	weak	(4)
CH1'	CH2' and CH2"	both equally strong	(5)
CH2'	CH3'	weak	(6)
CH2"	CH3'	nil	(7)
CH3'	CH4'	weak	(8)

Clearly both G and C residues have sugar conformation as 01'-endo in the complexed state. The relative intensities in NOE spectra are:

GH8 ... GH2'	strong	(9)
GH8 ... GH2"	nil	(10)
GH1'... GH2'	nil	(11)
GH1'... GH2"	strong	(12)
GH2'...GH3'	nil	(13)
GH2"...GH3'	nil	(14)
CH6 ...CH2'	strong	(15)
CH6 ... CH2"	nil	(16)
CH1'... CH2'	nil	(17)
CH1'... CH2"	nil	(18)
CH2'... CH3'		(19)
CH2"... CH3'		(20)

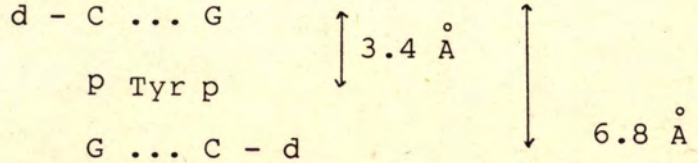
The structures are thus right-handed DNA with anti glyco-

sidic bond rotation for both the residues. Hence the conformation of d-CpG has not changed on complexation to Lys-Tyr-Lys.

The additional NOE's between protons those are not J-correlated, that is, between peptide and d-CpG protons from Figure 3.18 are:

NOE's pairs of protons	Distance from model	
CH6 ... Tyr H2, H6	(3.6 - 3.8 Å)	(21)
CH6 ... Tyr H3, H5	(4.0 Å)	(22)
GH3'... Tyr H2, H6	-	(23)
Tyr C^{α}... Tyr H2, H6	(2.0 Å)	(24)
Tyr C^{β}... Tyr H2, H6	(2.5 Å)	(25)
CH4' ... Tyr $C^{\beta}H$	(3.4 Å)	(26)
CH5' ... Tyr $C^{\beta}H$	(1.2 - 1.4 Å)	(27)

Thus NOE's are experimentally seen due to the fact that these pairs of protons are near to each other in space in the complex. The existence of NOE's no. (21) and (22) show that Tyr ring stacks on cytosine base. We have constructed dreiding stereo model of the complex by intercalating Tyr aromatic ring within the double helical d-CpG dinucleotide as shown below:



A front view of the model is shown in Figure 3.20. In order to change the distance between C - G and G - C base pair from 3.4 to 6.8 Å the backbone torsional angles are changed from standard B–DNA [7] geometry as follows:

Torsional angle	χ_{CN}	Φ'	ω'	ω	Φ	Ψ
B–DNA	-98°	155°	-96°	-46°	-147°	36°
d-CpG+Lys-Tyr-Lys model	-98°	-100°	210°	-190°	-140°	170°

The sugar conformations have not been changed on intercalating Tyr. The position of Tyr aromatic ring in a plane parallel to base pair has been adjusted in the model so as to result in a change in chemical shift on intercalation that is observed experimentally (Table 3.5) and is in conformity with the magnetic anisotropic effects due to the ring currents, as calculated by Pullman and Coworkers [47]. The distance between those pairs of protons for which NOE's are seen experimentally that is no. (21) to (27), in the model of complex, are given in paranenthesis along with NOE's. It is clearly seen that the distance data of the model complex agrees qualitatively with the observed NOE's. There is no direct evidence to suggest that the geometry shown in Figure 3.20 is the only one that could exist in solution.

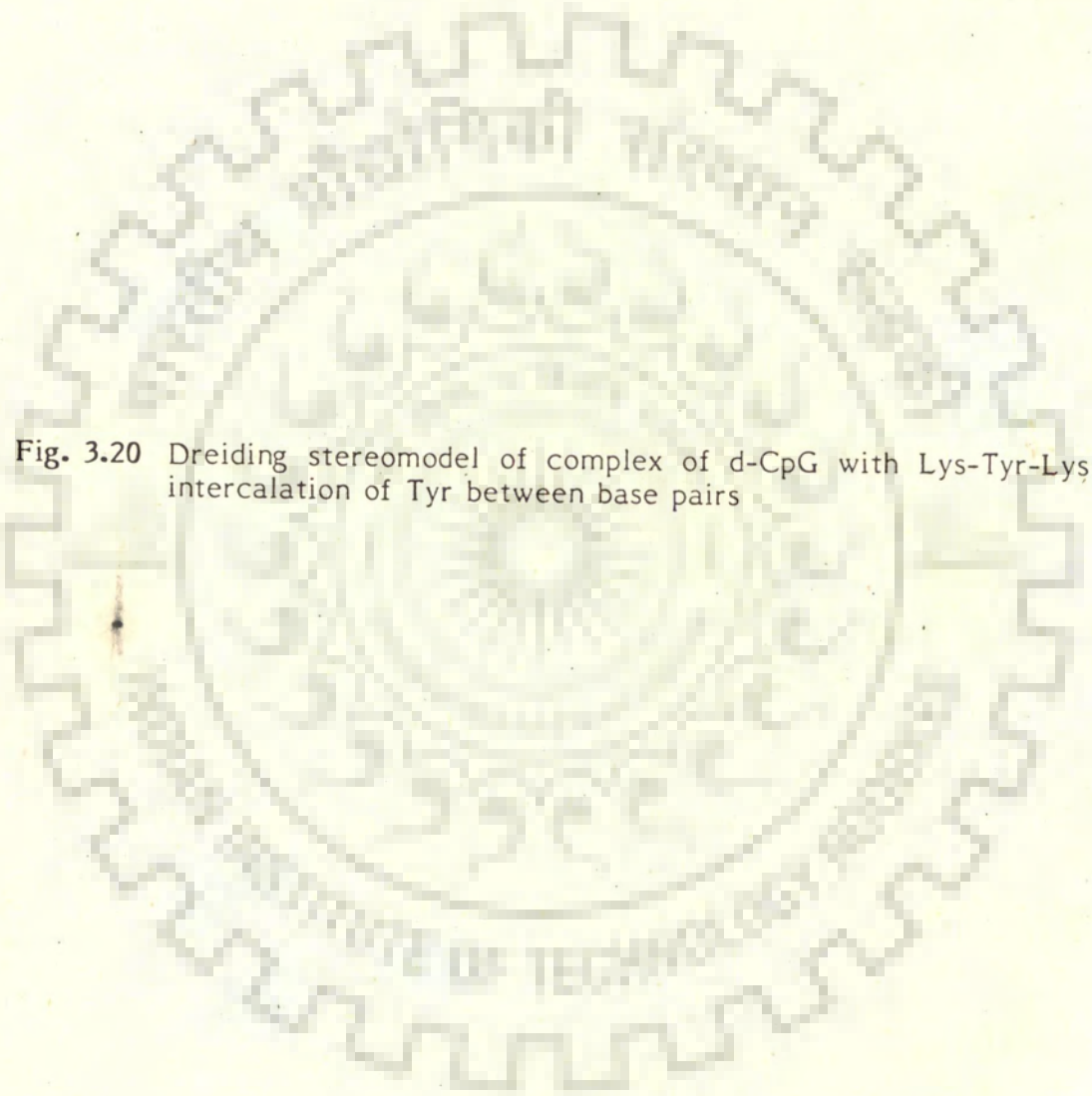
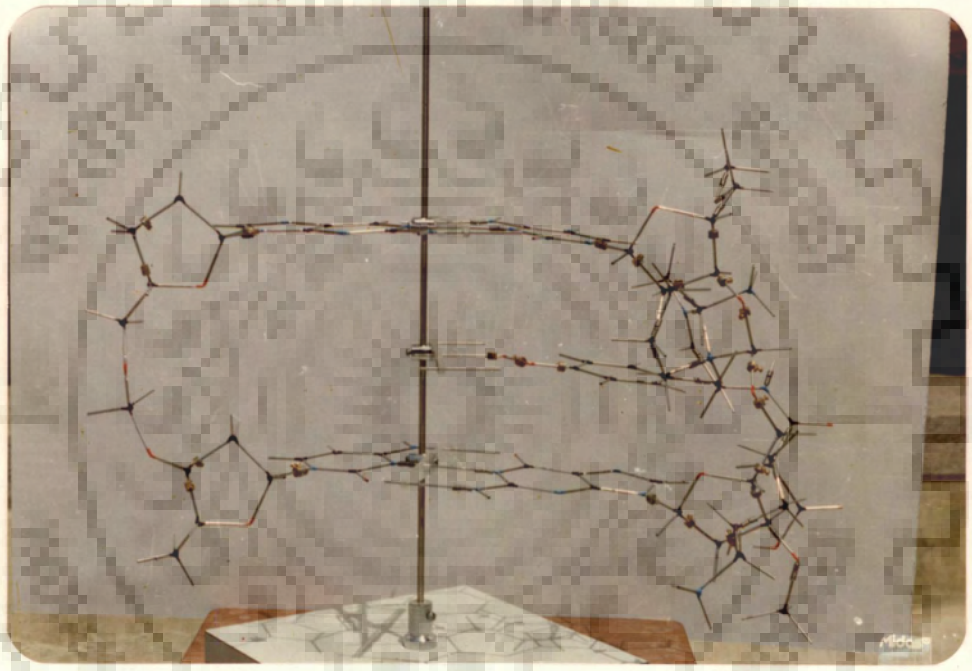


Fig. 3.20 Dreiding stereomodel of complex of d-CpG with Lys-Tyr-Lys showing intercalation of Tyr between base pairs



However, the model building studies do show that a specific geometry of complex can exist which is in agreement with the experimental NMR results, that is $\Delta\delta$, the upfield shift of Tyr ring, proton resonances, change in sugar conformation and interproton distances.

BINDING OF d-CpG TO Lys-Phe-Lys

Figure 3.21 and Table 3.6 show the results of proton NMR spectra of 5.0 mM d-CpG and its mixture with 2.5 mM Lys-Phe-Lys at 297 K. The spectral assignments have been made with the methods used earlier. The Phe ring protons shift upfield on binding of d-CpG to Lys-Phe-Lys which could be due to stacking of Phe with base pairs. The changes in chemical shift of GH8, CH6 and CH5 protons are relatively small. However, the large shift in Lys C^{ϵ} proton ~0.06-0.07 ppm indicates electrostatic interaction with negatively charged d-CpG.

Figures 3.22, 3.22a and 3.23 are the COSY and NOESY spectrum of the same sample solution as above at 297 K. The results on relative intensities of cross peaks in COSY spectra (see Figure 3.24) are:

GH1' ... GH2' and GH2"	both equally strong (not shown in fig.)	(1)
GH2' ... GH3'	strong	(2)

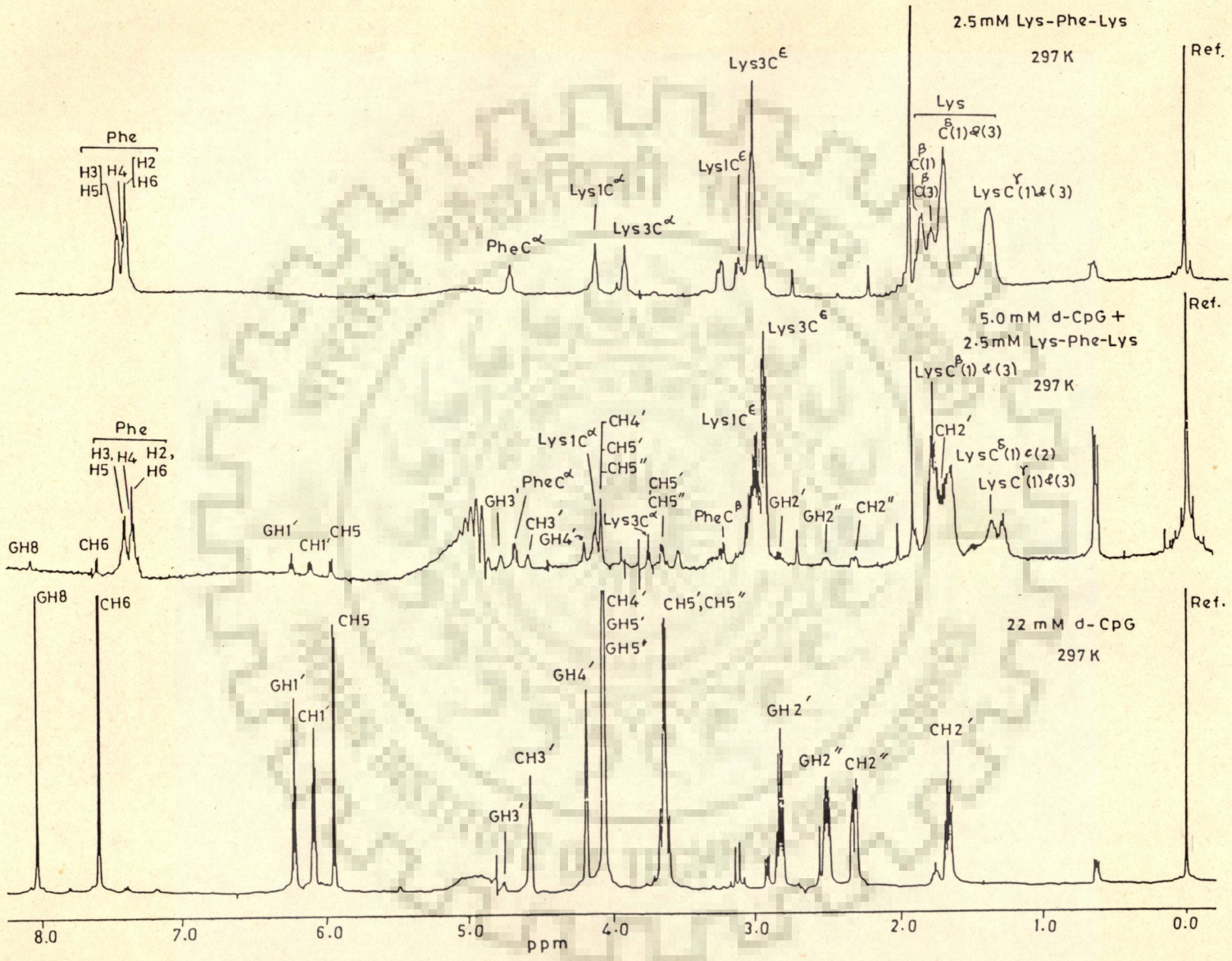


Fig. 3.21 500 MHz proton NMR spectra of 22 mM d-CpG, 2.5 mM Lys-Phe-Lys and a mixture of 5.0 mM d-CpG and 2.5 mM Lys-Phe-Lys in D_2O solution (pH=7.0) containing 0.125 mM EDTA at 297 K.

Table 3.6 : Changes in chemical shift, $\Delta\delta$ in ppm, of d-CpG and peptide protons on binding of 5.0 mM d-CpG to 2.5 mM Lys-Phe-Lys at 297 K.

SYSTEM	GH8	CH6	CH5	GH1'	CH1'	Phe H3,H5	Phe H4,H2,H6
d-CpG / Lys-Phe-Lys	8.032	7.578	5.945	6.216	6.080	7.377	7.319
d-CpG + Lys-Phe-Lys	8.050	7.567	5.949	6.218	6.094	7.370	7.310
$\Delta\delta$(ppm)*	-0.018	0.011	-0.004	-0.002	-0.014	0.007	0.009

SYSTEM	Phe C^{α}-H	Phe C^{β}-H	Lys C^{α}-H (1)	Lys C^{α}-H (3)	Lys C^{β}-H (1)	Lys C^{β}-H (3)	Lys C^{γ}-H (1)&(3)	Lys C^{δ}-H (1)&(3)	Lys C^{ε}-H (1)	Lys C^{ε}-H (3)
Lys-Phe-Lys	4.668	3.209	4.078	3.875	1.818	1.757	1.349	1.667	3.071	2.975
d-CpG + Lys-Phe-Lys	4.662	3.207	4.102	3.929	1.792	1.759	1.354	1.662	3.003	2.910
$\Delta\delta$ (ppm)*	0.006	0.002	-0.024	-0.054	0.026	-0.002	-0.005	0.005	0.068	0.065

*Upfield shifts, $\Delta\delta$ in ppm, are taken with positive sign.

Table 3.6 : Changes in chemical shift, $\Delta\delta$ in ppm, of d-CpG and peptide protons on binding of 5.0 mM d-CpG to 2.5 mM Lys-Phe-Lys at 297 K.

SYSTEM	GH8	CH6	CH5	GH1'	CH1'	Phe H3,H5	Phe H4,H2,H6
d-CpG / Lys-Phe-Lys	8.032	7.578	5.945	6.216	6.080	7.377	7.319
d-CpG + Lys-Phe-Lys	8.050	7.567	5.949	6.218	6.094	7.370	7.310
$\Delta\delta$(ppm)*	-0.018	0.011	-0.004	-0.002	-0.014	0.007	0.009

SYSTEM	Phe C^{α}-H	Phe C^{β}-H	Lys C^{α}-H		Lys C^{β}-H		Lys C^{γ}-H	Lys C^{δ}-H	Lys C^{ε}-H	
			(1)	(3)	(1)	(3)	(1)&(3)	(1)&(3)	(1)	(3)
Lys-Phe-Lys	4.668	3.209	4.078	3.875	1.818	1.757	1.349	1.667	3.071	2.975
d-CpG + Lys-Phe-Lys	4.662	3.207	4.102	3.929	1.792	1.759	1.354	1.662	3.003	2.910
$\Delta\delta$ (ppm)*	0.006	0.002	-0.024	-0.054	0.026	-0.002	-0.005	0.005	0.068	0.065

*Upfield shifts, $\Delta\delta$ in ppm, are taken with positive sign.

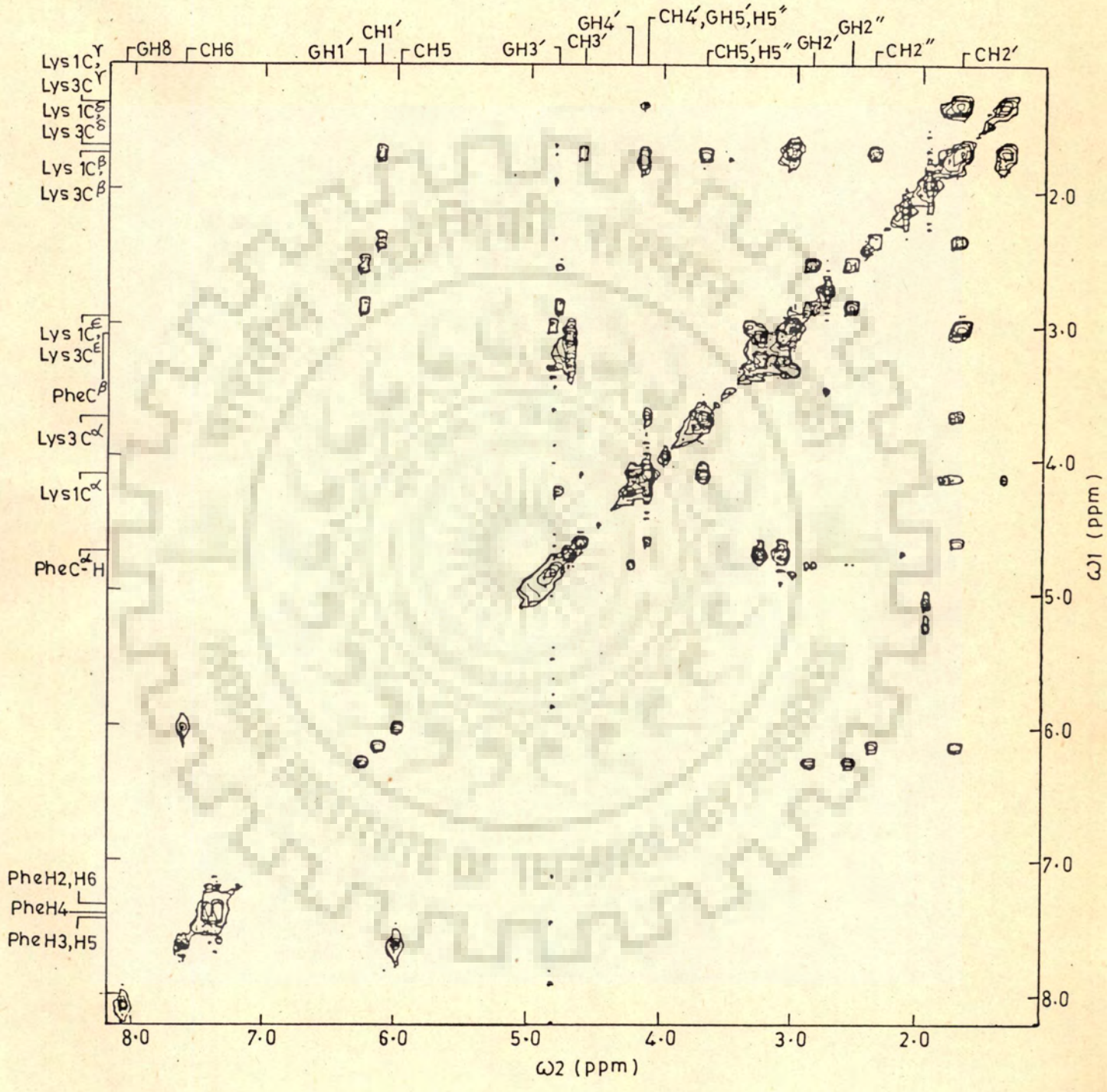


Fig. 3.22 500 MHz COSY spectrum of a mixture of 5.0 mM d-CpG and 2.5 mM Lys-Phe-Lys at 297 K for the sample used in Fig.3.21

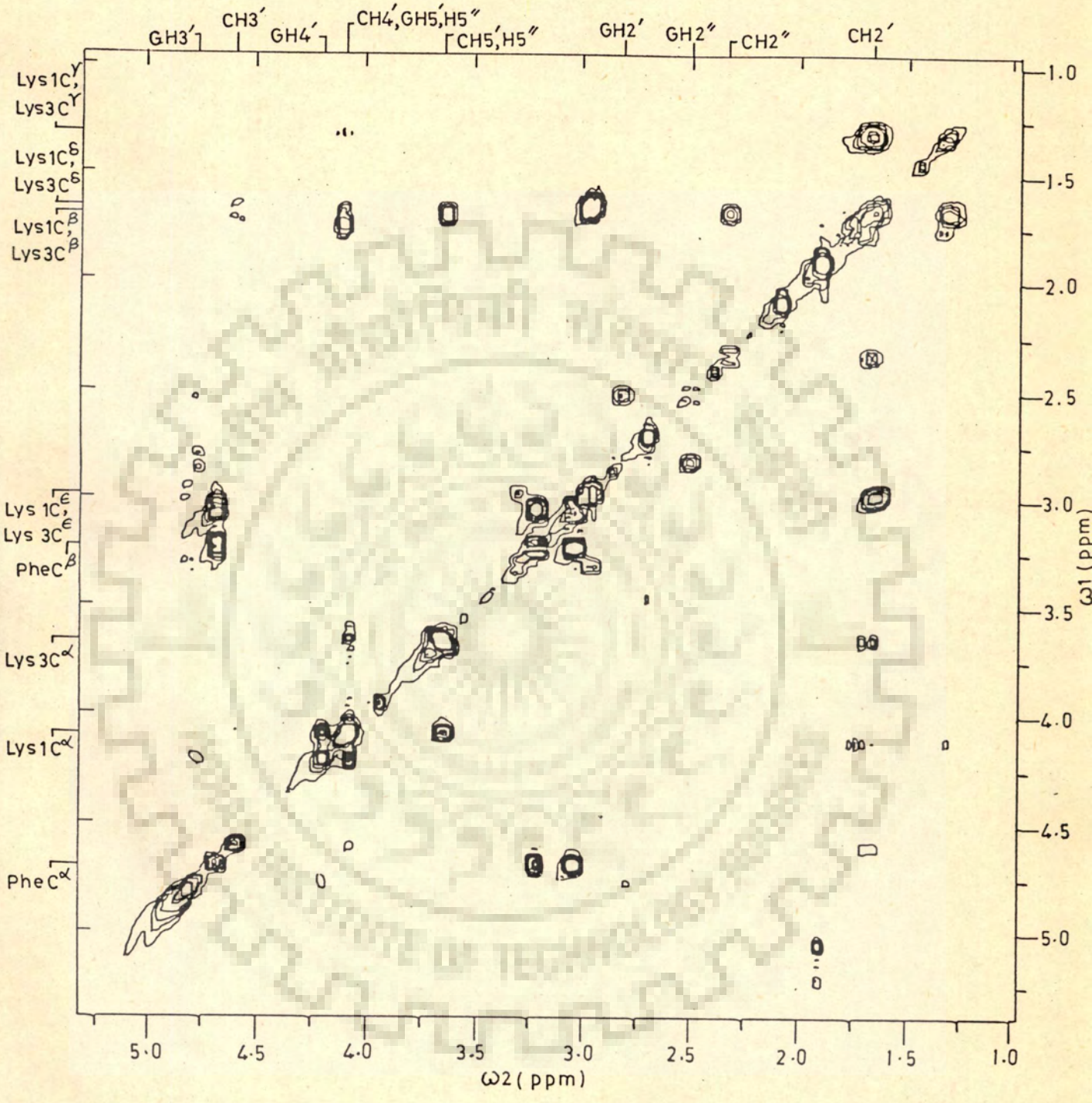


Fig. 3.22 (a) Portion of COSY spectrum of Fig.3.22 in an expanded scale.

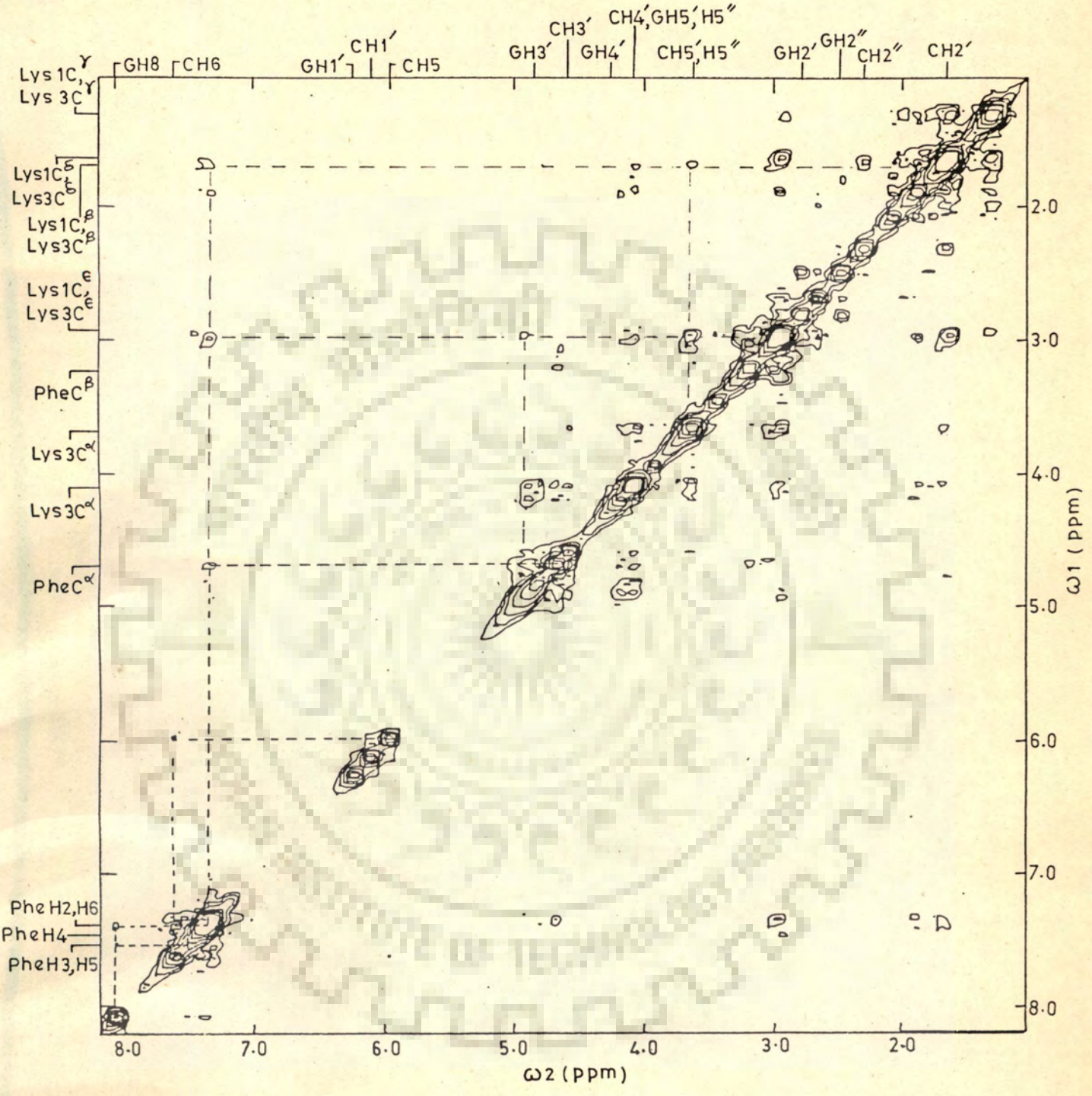


Fig. 3.23 500 MHz NOESY spectrum of a mixture of 5.0 mM d-CpG and 2.5 mM Lys-Phe-Lys at 297 K for the sample used in Fig.3.21

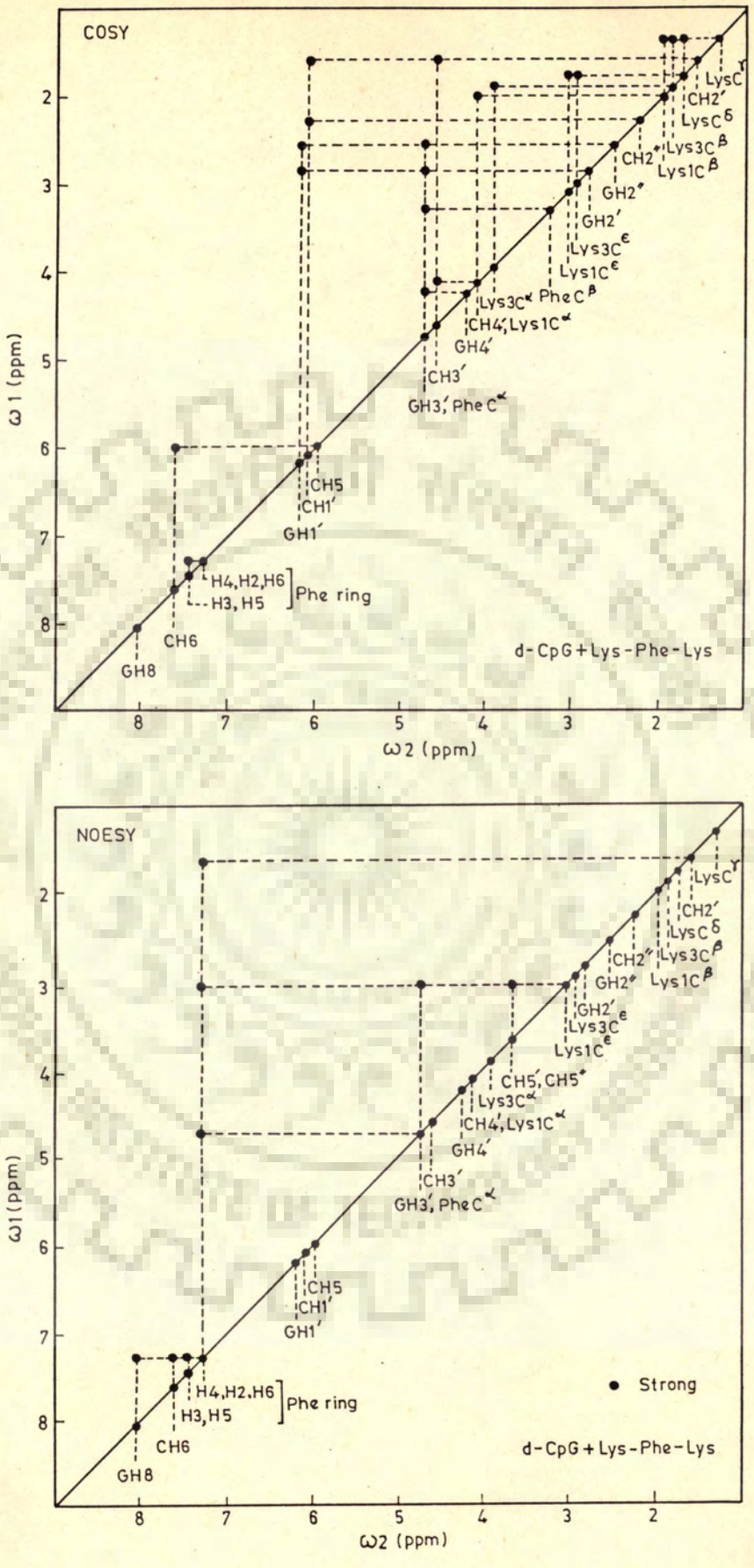


Fig. 3.24 Schematic representation of the results of COSY and NOESY spectra of a mixture of d-CpG with Lys-Phe-Lys shown in Fig. 3.22 & 23.

GH2" ... GH3'	strong	(3)
GH3' ... GH4'	strong	(4)
CH1' ... CH2' and CH2"	both equally strong	(5)
CH2' ... CH3'	strong	(6)
CH2" ... CH3'	Nil	(7)
CH3' ... CH4'	strong	(8)

Clearly as per strategies discussed earlier both G and C are in Ol-endo sugar conformations in the complex of d-CpG with Lys-Phe-Lys. The intensities in NOE spectra are rather low so that many of the expected NOE's are not seen. However, several NOE's between different protons of peptide and nucleotide are seen. These are:

GH8 ... Phe ring protons	(9)
CH6 ... Phe ring protons	(10)
Phe C^{α} ... Phe ring protons	(11)
Lys C^{ε} ... Phe ring protons	(12)
CH2' ... Phe ring protons	(13)
GH3' ... Lys C^{ε}	(14)
CH5', CH5"... Lys C^{ε}	(15)

The NOE's no. (9) and (10) demonstrate that Phe ring intercalates betwen C-G base pairs of d-CpG whereas others define a specific geometry of complex.

CHAPTER-4

d-GpCpGpC AND ITS BINDING TO TRIPEPTIDE Lys-Tyr-Lys

500 MHz proton NMR spectra of deoxytetranucleotide d-GpCpGpC and its mixture with tripeptide Lys-Tyr-Lys as a function of temperature in the range 285-355K were studied. 2D-COSY and NOESY spectra were taken to find the conformation of d-GpCpGpC alone and its complex with tripeptide.

d-GpCpGpC

A. Assignments

Figure 4.1 shows a typical spectra of d-GpCpGpC in D_2O as a function of temperture. Nonexchangeable base protons are easily identified from their chemical shift values. For instance GH8 as expected, appear at about 8.0 ppm, CH6 around 7.5 ppm and CH5 near 5.5 ppm. A group of H1' sugar protons appears around 6.0 ppm and H2', H2" protons at 2-3 ppm. The exact sequential assignment has, however, been made possible only on the basis of 2D-COSY and NOESY spectra, discussed in next paragraph. The bases in d-GpCpGpC have been numbered as follows:

d - G p C p G p C
1 2 3 4
ext. int. int. ext.

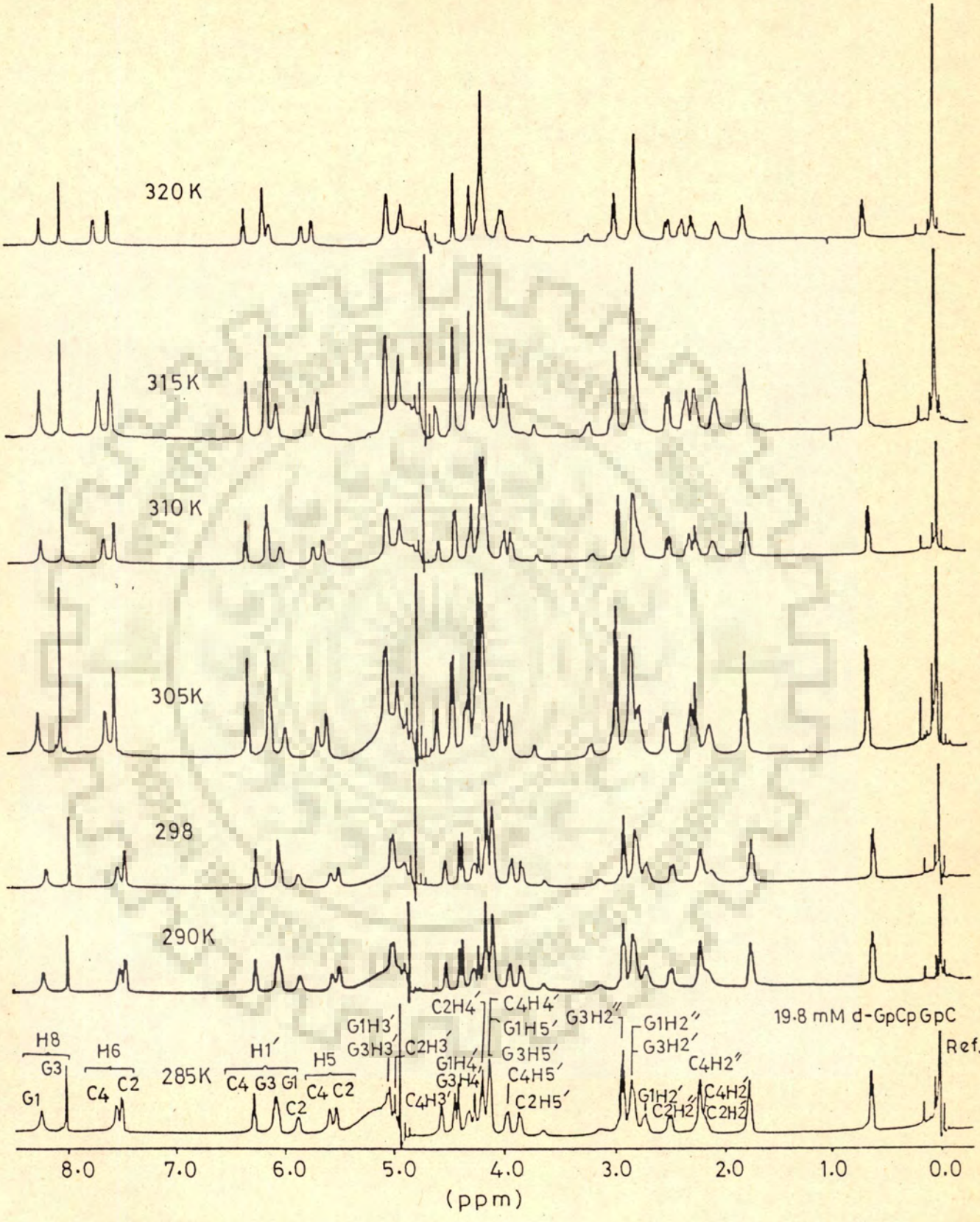


Fig. 4.1 500 MHz proton NMR spectra of deoxytetranucleotide 19.8 mM d-GpCpGpC in D_2O (pH=7.0) solution containing 0.25 mM EDTA at indicated temperatures. Ref. DSS.

The observed chemical shifts as a function of temperature are given in Table 4.1 and Fig. 4.2. The large downfield shift in ring protons of base CH6, CH5 is due to predominant destacking of bases with increase in temperature. The changes in chemical shift values with temperature are ascribed to opening of mini-double helix of d-GpCpGpC into single-strands. The $T_{\frac{1}{2}}$ value from Fig. 4.2 is about 313-315K. Figs. 4.3 and 4.4 show 500 MHz two-dimensional COSY and NOESY spectra of d-GpCpGpC at 297 K while Figs. 4.3a, b, c and 4.4a, b, c show portions of these spectra expanded to show specific connectivities. The sample solution used is the same as that in Fig. 4.1. The sequential assignments have been made using strategies discussed earlier. Two triplets of H1' protons appear at 6.26 and 5.88 ppm while the other two H1' proton triplets overlap at about 6.05 ppm at 297 K. Since these are all J-coupled to corresponding H2' and H2" the four sets of H2' and H2" are easily identified in the COSY plot. For three sets of sugar resonances, both H2' as well as H2" give connectivity with H3' while for one of the sugars only H2' gives connectivity with H3' in COSY plots. The various COSY and NOE connectivities are reproduced conveniently in Figs. 4.5 and 4.6 which are schematic representation derived from the experimental results in Figs. 4.3 and 4.4. Each H3' proton is coupled to H4' and each H4' is further coupled to H5' and H5" so that four sets of sugar

Table 4.1 : Chemical shift values of various protons of 500 MHz for 19.8 mM d-GpCpGpC in D_2O (pH=7.1) in temperature range 285K-320K

Temp. (K)	G1H8	G3H8	C2H6	C4H6	G1H1'	G3H1'	C2H1'	C4H1'
285	8.235	8.001	7.475	7.522	6.054	6.067	5.857	6.260
290	8.224	7.992	7.470	7.520	6.049	6.049	5.862	6.258
297	8.210	7.983	7.470	7.527	6.046	6.060	5.885	6.259
305	8.191	7.972	7.473	7.555	6.042	6.070	5.918	6.258
310	8.176	7.965	7.481	7.757	6.063	6.063	5.949	6.257
315	8.160	7.958	7.491	7.595	6.069	6.069	5.976	6.255
320	8.145	7.954	7.507	7.638	6.086	6.086	5.020	6.255

Temp. (K)	C2H5	C4H5	C4H2'/ C2H2'	C4H2"	C2H2"	G1H2'	G1H2"/ G3H2'	G3H2"
285	5.510	5.572	2.219	2.219	2.493	2.708	2.831	2.895
290	5.508	5.570	2.170	2.217	2.486	2.703	2.820	2.894
297	5.525	5.594	2.128	2.225	2.478	2.710	2.806	2.893
305	5.552	5.630	2.085	2.218	2.471	2.718	2.783	2.890
310	5.570	5.659	2.058	2.213	2.455	2.705	2.768	2.889
315	5.593	5.682	2.031	2.218	2.441	2.737	2.737	2.888
320	5.641	5.735	2.010	2.223	2.435	2.719	2.719	2.885

Temp. (K)	G1H3'/ G3H3'	C4H3'	G1H4'	G3H4'	C2H4'	C4H4',G1H5'/ G3H5'	C2H5'	C4H5'
285	5.024	4.549	4.424	4.391	4.170	4.113	3.858	3.948
290	5.020	4.540	4.417	4.385	4.167	4.105	3.857	3.943
297	5.014	4.538	4.410	4.381	4.163	4.112	3.856	3.951
305	5.004	4.535	4.397	4.374	4.153	4.110	3.867	3.948
310	5.988	4.531	4.388	4.370	4.142	4.112	3.887	3.935
315	5.980	4.532	4.370	4.370	4.115	4.115	3.893	3.934
320	4.975	-	4.365	4.365	-	4.110	3.912	3.935

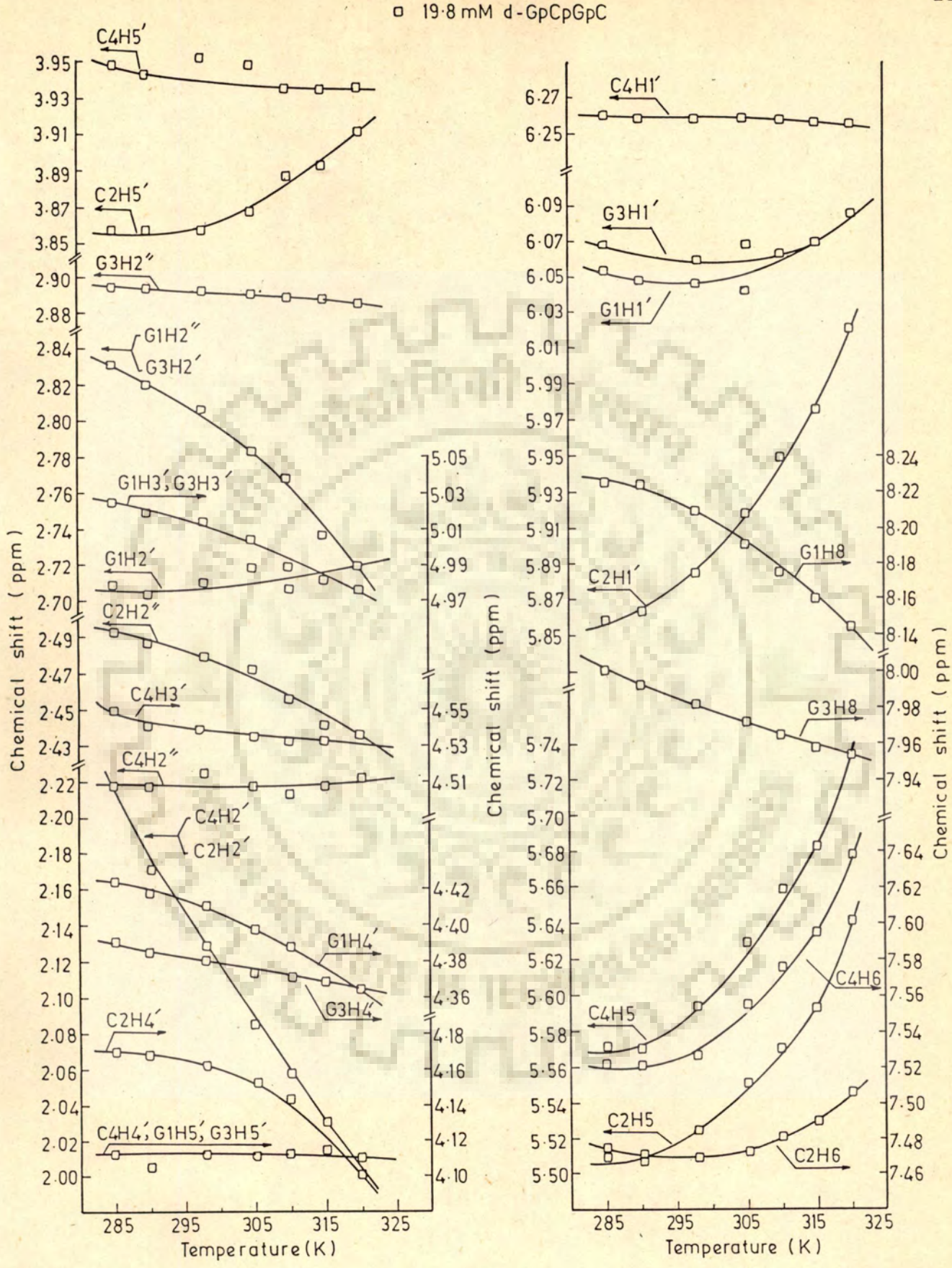


Fig. 4.2 Chemical shift of various protons of d-GpCpGpC as a function of temperature taken from spectra of Fig. 4.1

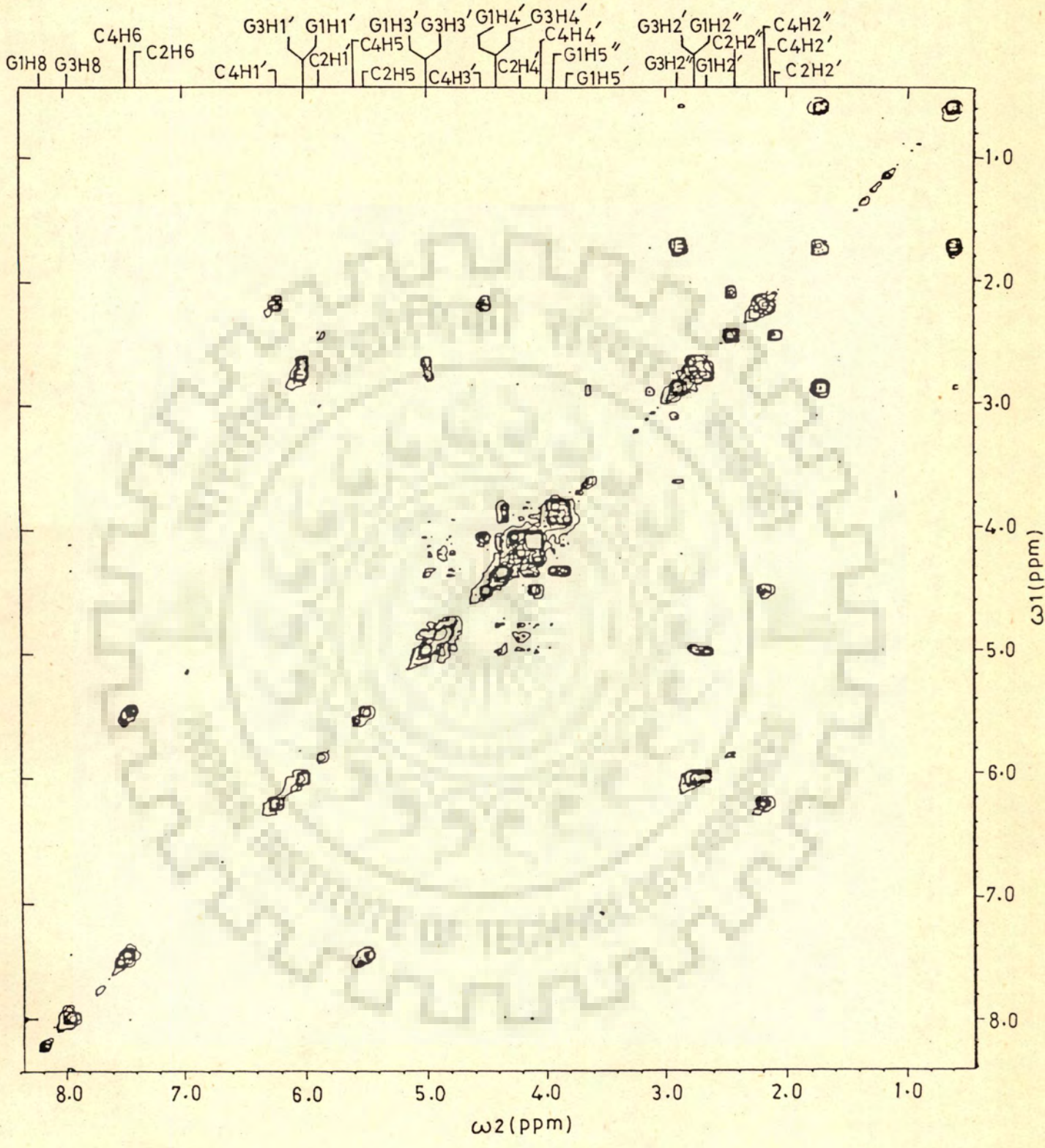
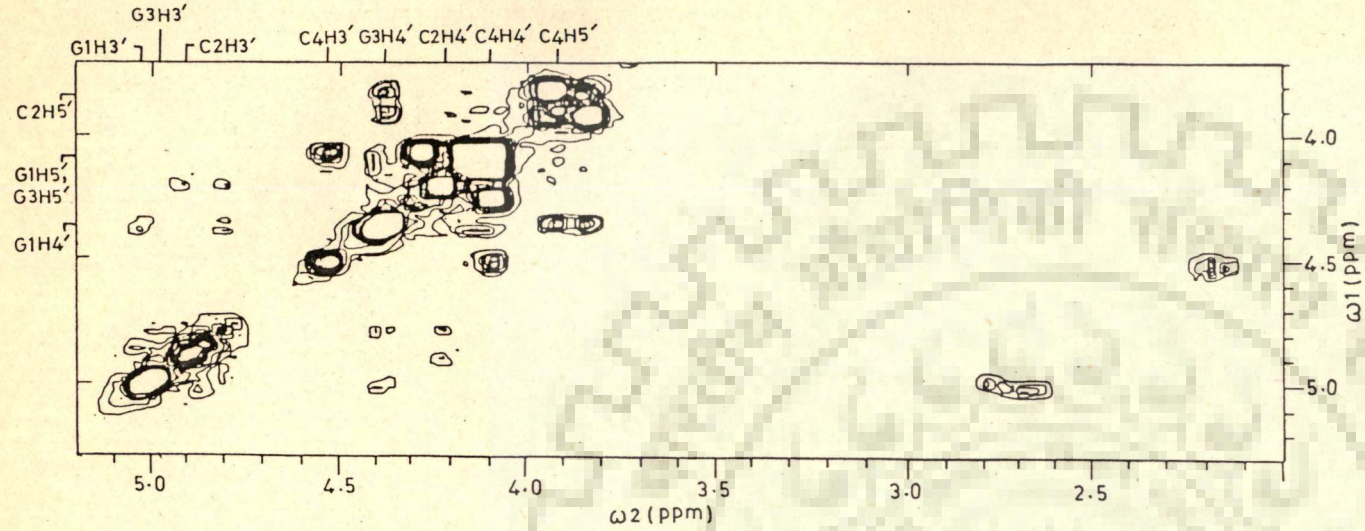
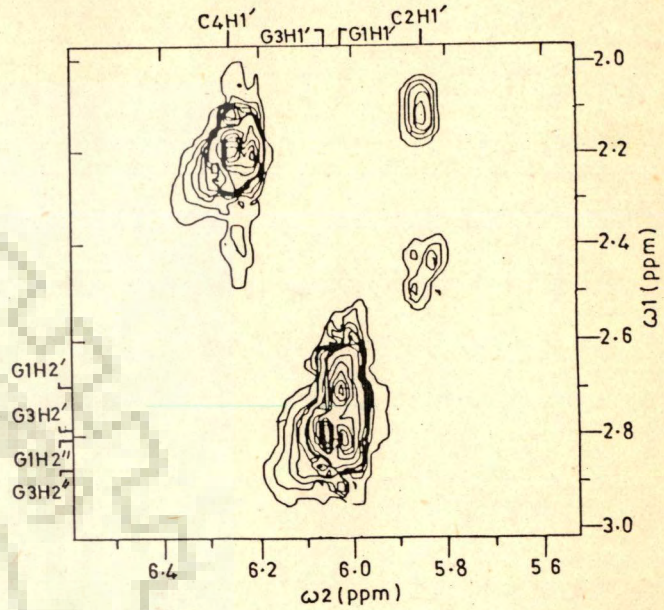


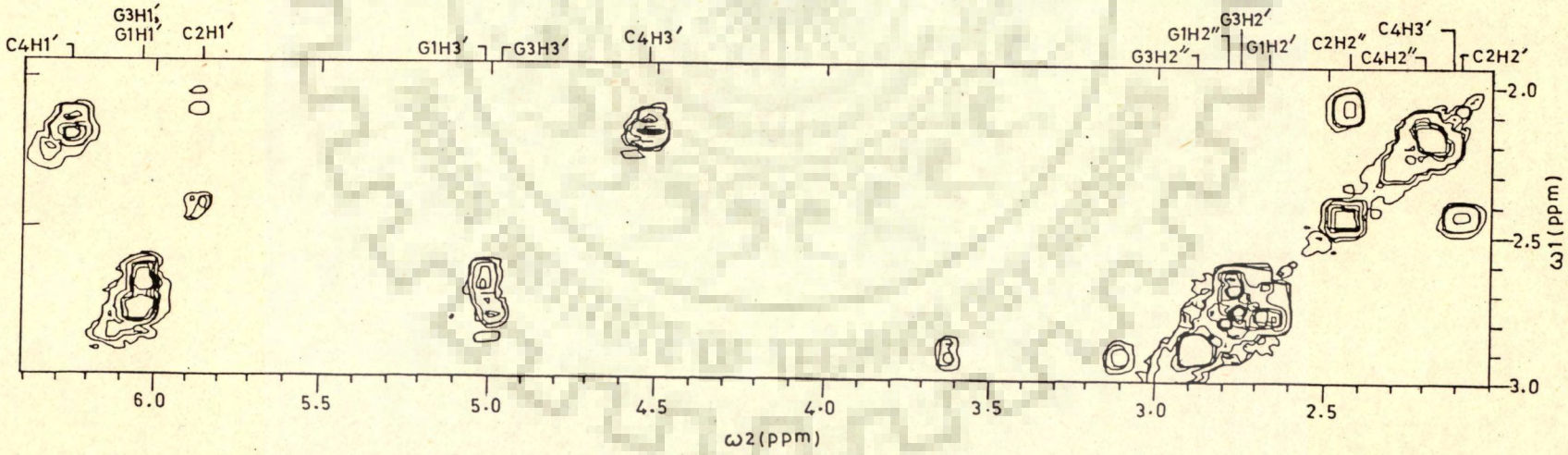
Fig. 4.3 500 MHz COSY spectrum of 19.8 mM d-GpCpGpC solution of in D_2O containing 0.25 mM EDTA (pH=7.0) at 297 K



(a)



(b)



(c)

Fig. 4.3 (a), (b), (c) Portions of COSY spectrum of Fig. 4.3. shown in expanded scales.

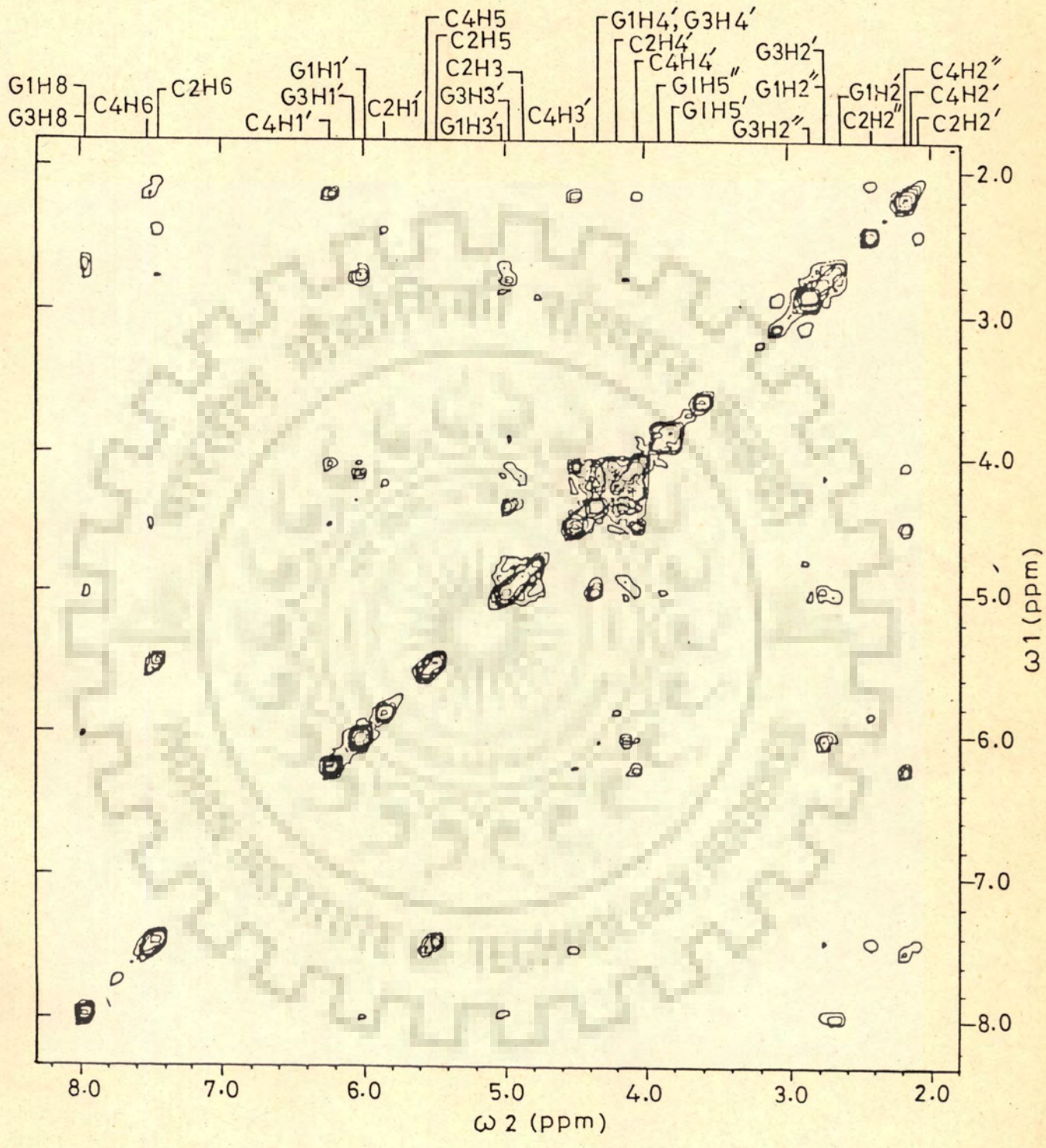
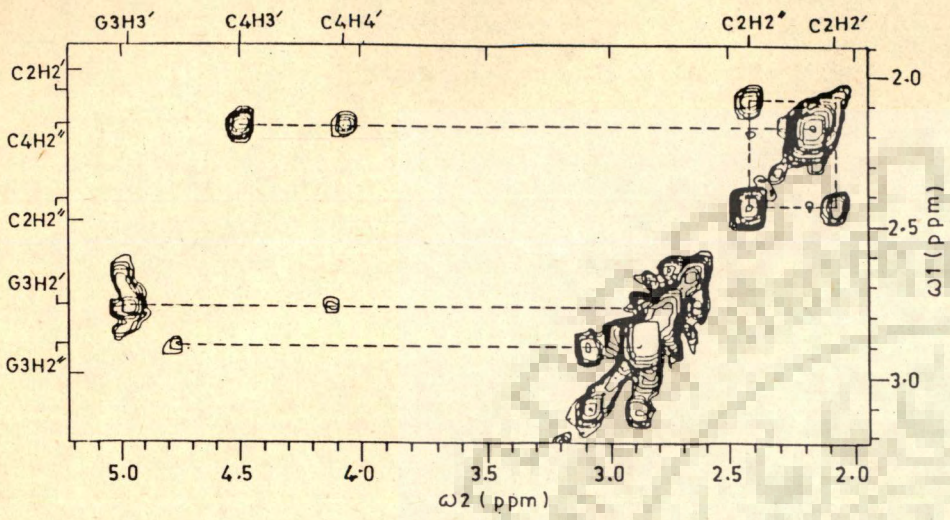
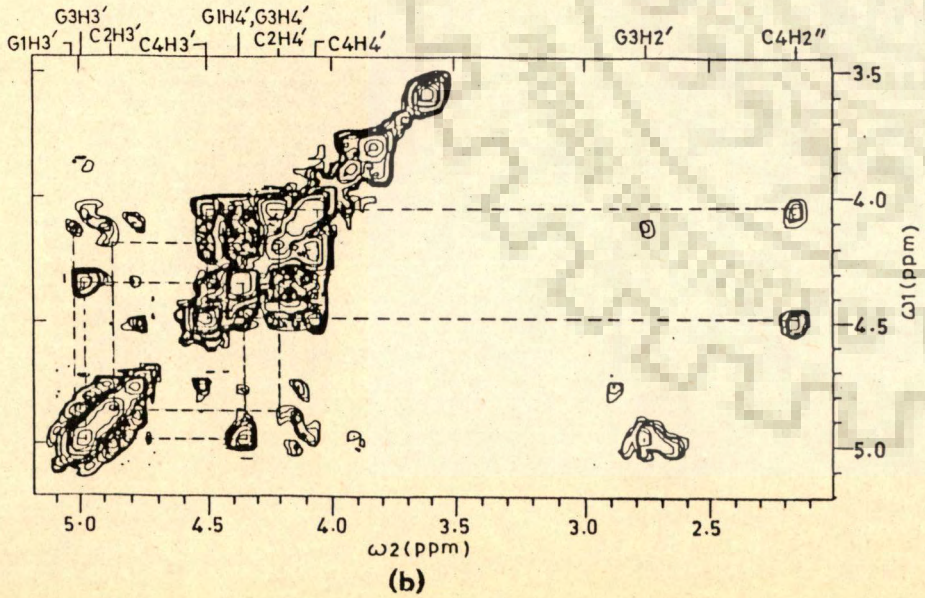


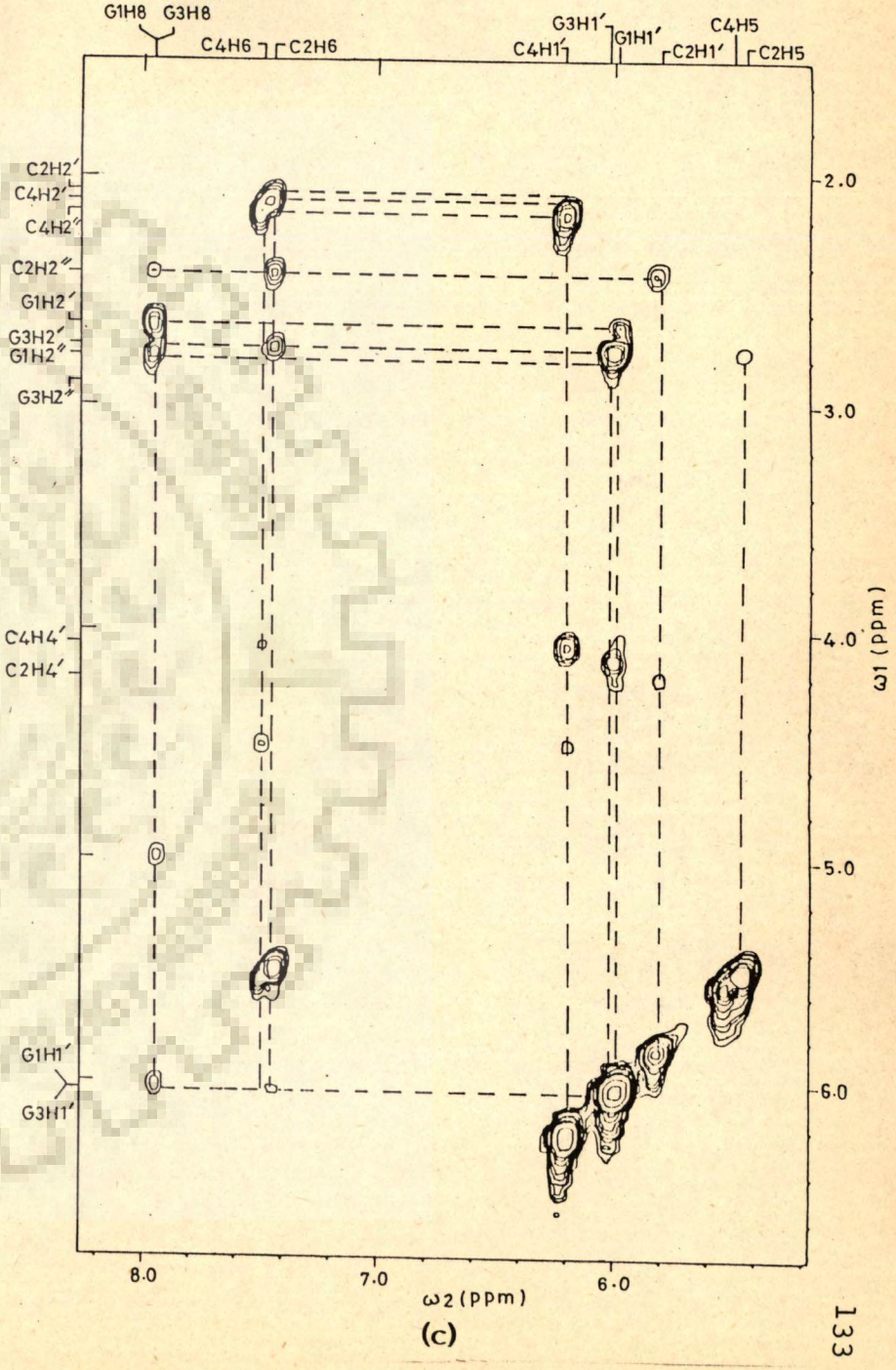
Fig. 4.4 500 MHz NOESY spectrum of 19.8 mM d-GpCpGpC solution in D_2O containing 0.25 mM EDTA (pH=7.0) at 297 K.



(a)



(b)



(c)

Fig. 4.4 **(a), (b), (c)** Portions of NOESY spectrum of Fig. 4.4 in expanded

resonances are easily identified.

Assignment of each set of sugar protons to specific base in d-GpCpGpC is done with the help of NOE connectivities. Within a nucleotide base ring proton GH8 or CH6 is close (closest) to the corresponding H2' and to H2" and H1' (intra-residue connectivities). Further in right-handed 5'-3' oligonucleotide sequence the base proton, GH8/CH6 is close to the H2" and H1' protons of the nucleotide preceeding it (inter-residue connectivities). Thus using the strategy (n, stands for nth residue in 5'-3' sequence)

$(GH8/CH6)_n \ldots (H2')_n, (H2")_n, (H1')_n$ intra-residue

$(GH8/CH6)_n \ldots (H2")_{n-1}, (H1')_{n-1}$ inter-residue

The following NOE connectivities are possible in d-GpCpGpC:

G1H8... G1H2'	intra-residue	(1)
G1H8... G1H2"	-do-	(2)
G1H8... G1H1'	-do-	(3)
C2H6...G1H1'	inter-residue	(4)
C2H6...G1H2"	-do-	(5)
C2H6...C2H2'	intra-residue	(6)
C2H6...C2H2"	-do-	(7)
C2H6...C2H1'	-do-	(8)

G3H8... C2H1'	inter-residue	(9)
G3H8... C2H2"	-do-	(10)
G3H8... G3H2'	intra-residue	(11)
G3H8... G3H2"	-do-	(12)
G3H8... G3H1'	-do-	(13)
C4H6... G3H1'	inter-residue	(14)
C4H6... G3H2"	-do-	(15)
C4H6... C4H2'	intra-residue	(16)
C4H6... C4H2"	-do-	(17)
C4H6... C4H1'	-do-	(18)

In Figs. 4.4 and 4.6, the appearance of intra-residue pairs of connectivities (6)-(7), (16)-(17), (1)-(2), (11)-(12) indicate which of the H2'-H2" pairs belong to which particular bases. The appearance of inter-residue connectivities (5), (10) and (15) clearly assign two GH8's to G1 and G3 and further two CH6's to C2 and C4, respectively. Having made these assignments the 4 triplets of H1', H3' H4', H5' and H5" get straightaway assigned to specific bases. The complete unambiguous assignment was thus made.

B. Conformation

To obtain sugar conformations it is noted (see Fig. 4.5) that each of the bases H1' is J-coupled to

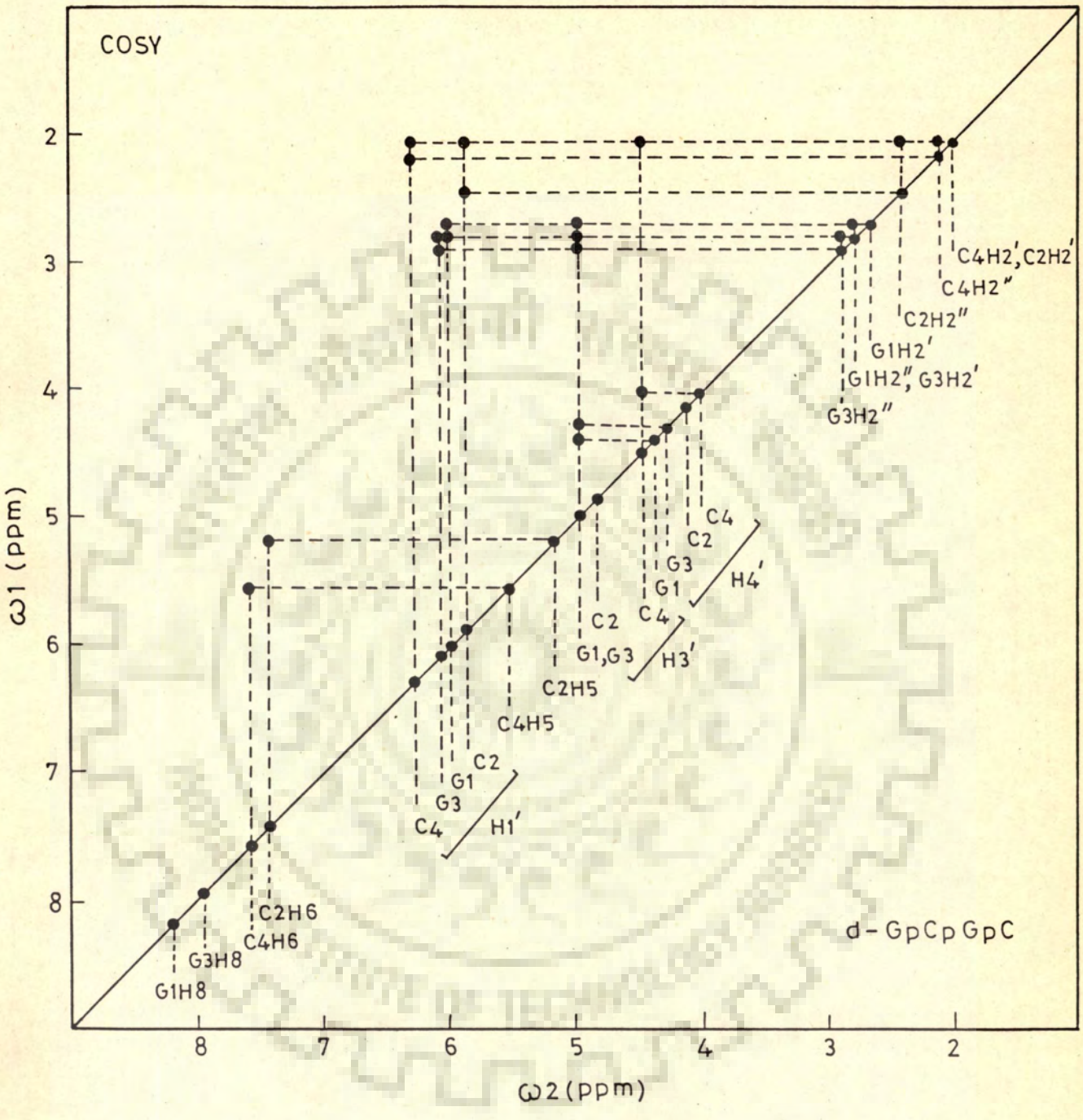


Fig. 4.5 Schematic representation of the results COSY spectra of d-GpCpGpC shown in Fig. 4.3.

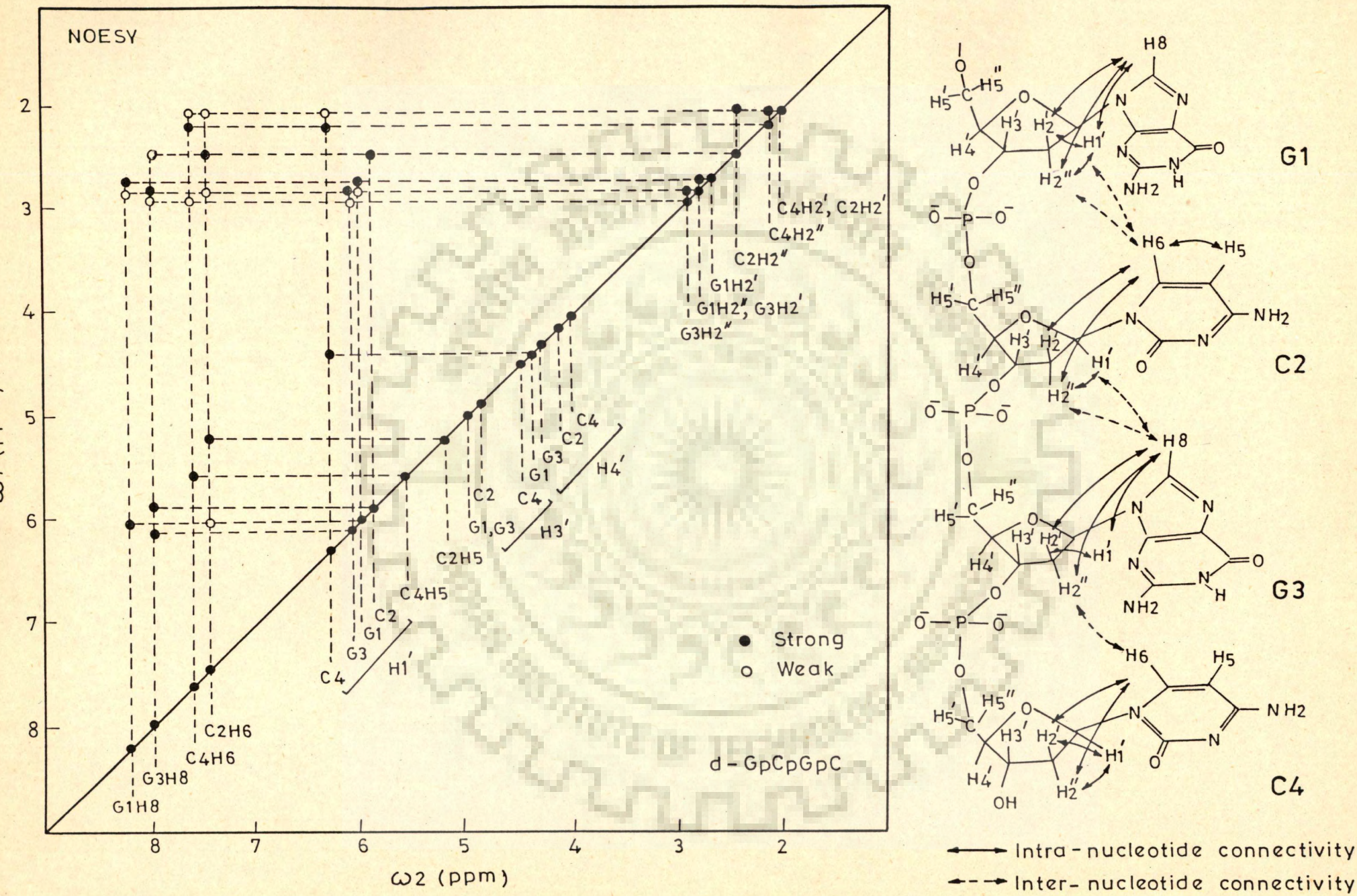


Fig. 4.6 Schematic representation of the results of NOESY spectra of d-GpCpGpC shown in Fig. 4.4, alongwith 5'-3' d-GpCpGpC oligonucleotide giving various intra - and inter-residue NOE connectivities.

its corresponding H2' and H2". H2' is coupled to H3' and H3' is coupled to H4'. Therefore, all sugars have O1'-endo conformation.

The relative intensities of NOE cross peaks are:

G1H8... G1H2'	intra-residue	strong	(1)
G1H8... G1H2"	-do-	weak	(2)
G1H8... G1H1'	-do-	strong	(3)
C2H6... G1H1'	inter-residue	weak	(4)
C2H6... G1H2"	-do-	weak	(5)
C2H6... C2H2'	intra-residue	weak	(6)
C2H6... C2H2"	-do-	strong	(7)
C2H6...C2H1'	-do-	nil	(8)
G3H8... C2H1'	inter-residue	strong	(9)
G3H8... C2H2"	-do-	weak	(10)
G3H8... G3H2'	intra-residue	strong	(11)
G3H8... G3H2"	-do-	weak	(12)
G3H8... G3H8'	-do-	strong	(13)
C4H6... G3H1'	inter-residue	not seen	(14)
C4H6... G3H2"	-do-	weak	(15)
C4H6... C4H2'	intra-residue	weak	(16)
C4H6... C4H2"	-do-	strong	(17)

C4H6... C4H1'	intra-residue	nil	(18)
G1H1'...G1H2'	-do-	strong	(19)
G1H1'...G1H2"	-do-	weak	(20)
C2H1'...C2H2'	-do-	nil	(21)
C2H1'...C2H2"	-do-	strong	(22)
G3H1'...G3H2'	-do-	strong	(23)
G3H1'...G3H2"	-do-	weak	(24)
C4H1'...C4H2'	-do-	weak	(25)
C4H1'...C4H2"	-do-	strong	(26)
C4H1'...G1H4'	inter-residue	strong	(27)

For G1 and G3 the base to H2' connectivity [(1) and (11)] is stronger than that of base to H2" connectivity [(2) and (12)] so that the glycosidic bond rotation χ_{CN} is inferred to be anti conformation. However for C2 and C4, the base to H2" connectivity [(7) and (17)] is stronger than that of base to H2' [(6) and (16)] so that χ_{CN} for these two bases is in high anti conformation.

The appearance of cross peak no. (27) that is C4H1' with G1H4' demonstrates that most of d-GpCpGpC at room temperature exists as a double-helix as

```
          1   2   3   4
d - G   C   G   C
    |   |   |   |
    C   G   C   G - d
          4   3   2   1
```

The existence of set of cross peaks no. (1), (6), (11), (16) that is, connectivity between GH8/CH6 to its corresponding H2' and a set of cross peak no. (5), (10), (15) that is, connectivity between $(GH8/CH6)_n$ to $(H2")_{n-1}$ indicates that the deoxyoligonucleotide exists as a right-handed B-DNA structure. Also the absence of any connectivity of $(GH8)_n$ with $(CH5")_{n-1}$ further shows that there is no left-handed structure existing is our sample solution.

Hence, it is concluded that d-GpCpGpC exists as a right-handed double helical DNA at 285 K and changes to single-strand with increase in temperature, yielding $T_{\frac{1}{2}} \sim 314$ K. The sugar conformations are O1'-endo for all bases whereas glycosidic bond rotation is anti for G1 and G3 but anti/high anti for C2 and C4 bases.

BINDING OF d-GpCpGpC TO Lys-Tyr-Lys

Figs. 4.7 and 4.8 show the results of binding of d-GpCpGpC to tripeptide Lys-Tyr-Lys at different temperatures. Changes in chemical shift of all protons on binding are also given in Table 4.2. The peptide proton assignments have been made as discussed in Chapter-3. These results clearly show that nucleotide protons exhibit much more pronounced melting curves than that in the absence of peptide. The binding of d-GpCpGpC to peptide has thus stabilised the double-helix so that the difference in

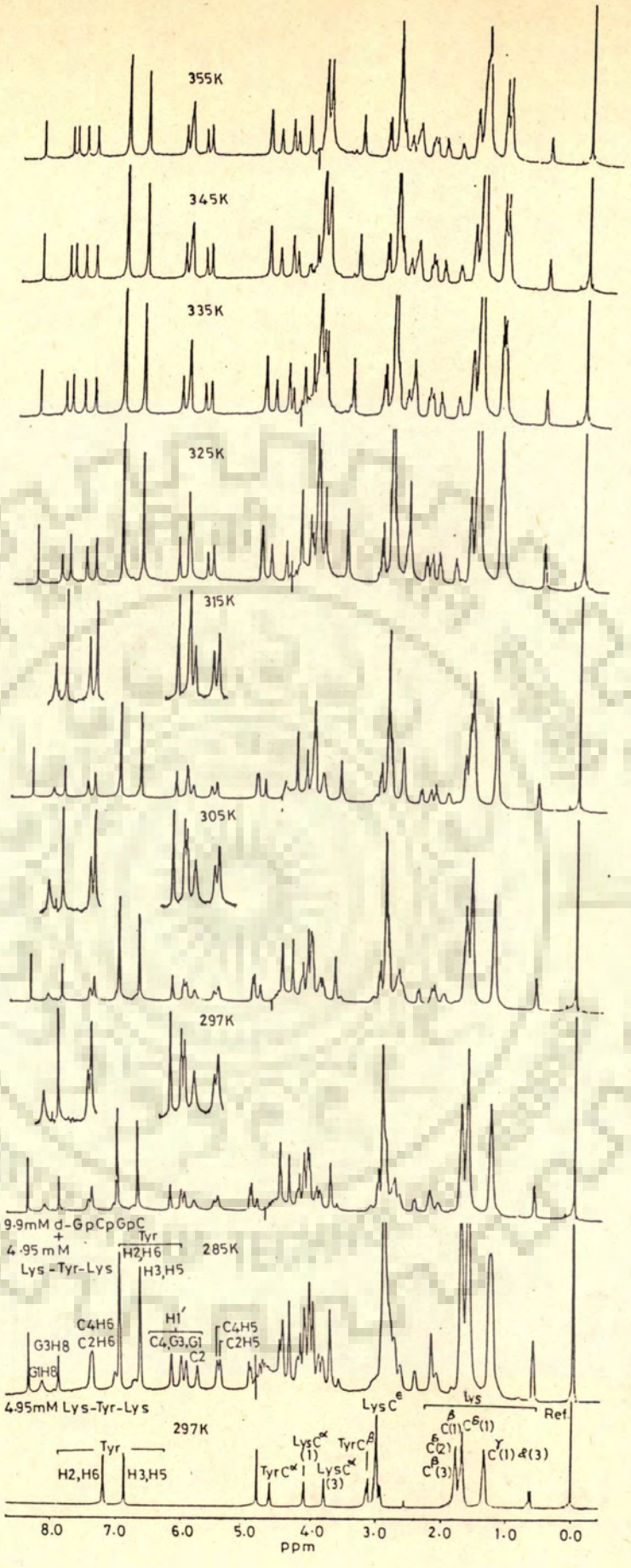


Fig. 4.7 500 MHz proton NMR spectra of a mixture of 9.9 mM d-GpCpGpC and 4.95 mM Lys-Tyr-Lys at different temperatures in D_2O (pH=7.1) containing 0.25 mM EDTA. For comparison 500 MHz spectrum of 4.95 mM Lys-Tyr-Lys in D_2O solution at 297 K is also shown. Ref. DSS.

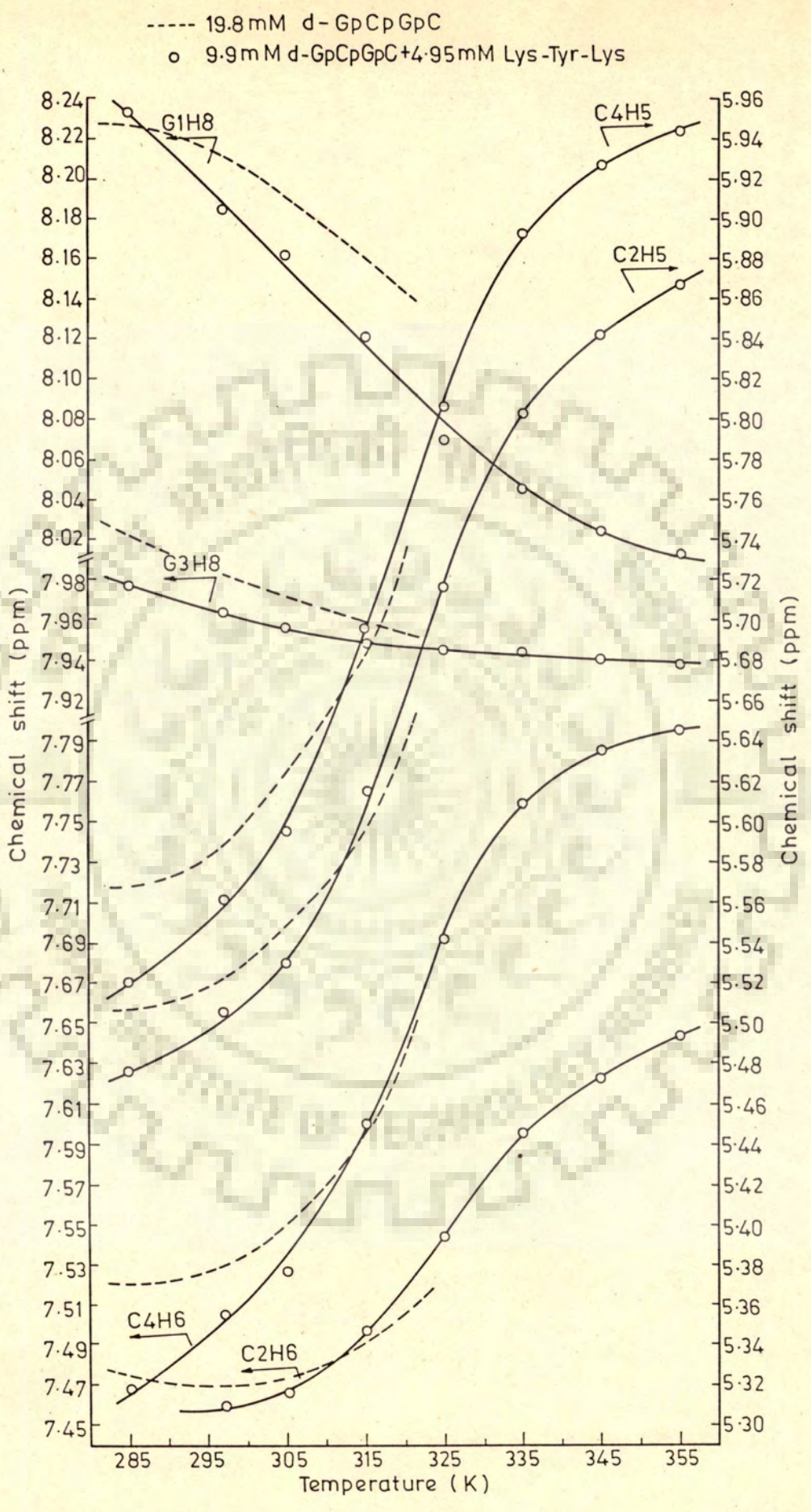


Fig. 4.8 Chemical shifts of some of the nucleotide and peptide protons as a function of temperature in uncomplexed nucleotide/peptide and the complexed state from data in Fig. 4.7.

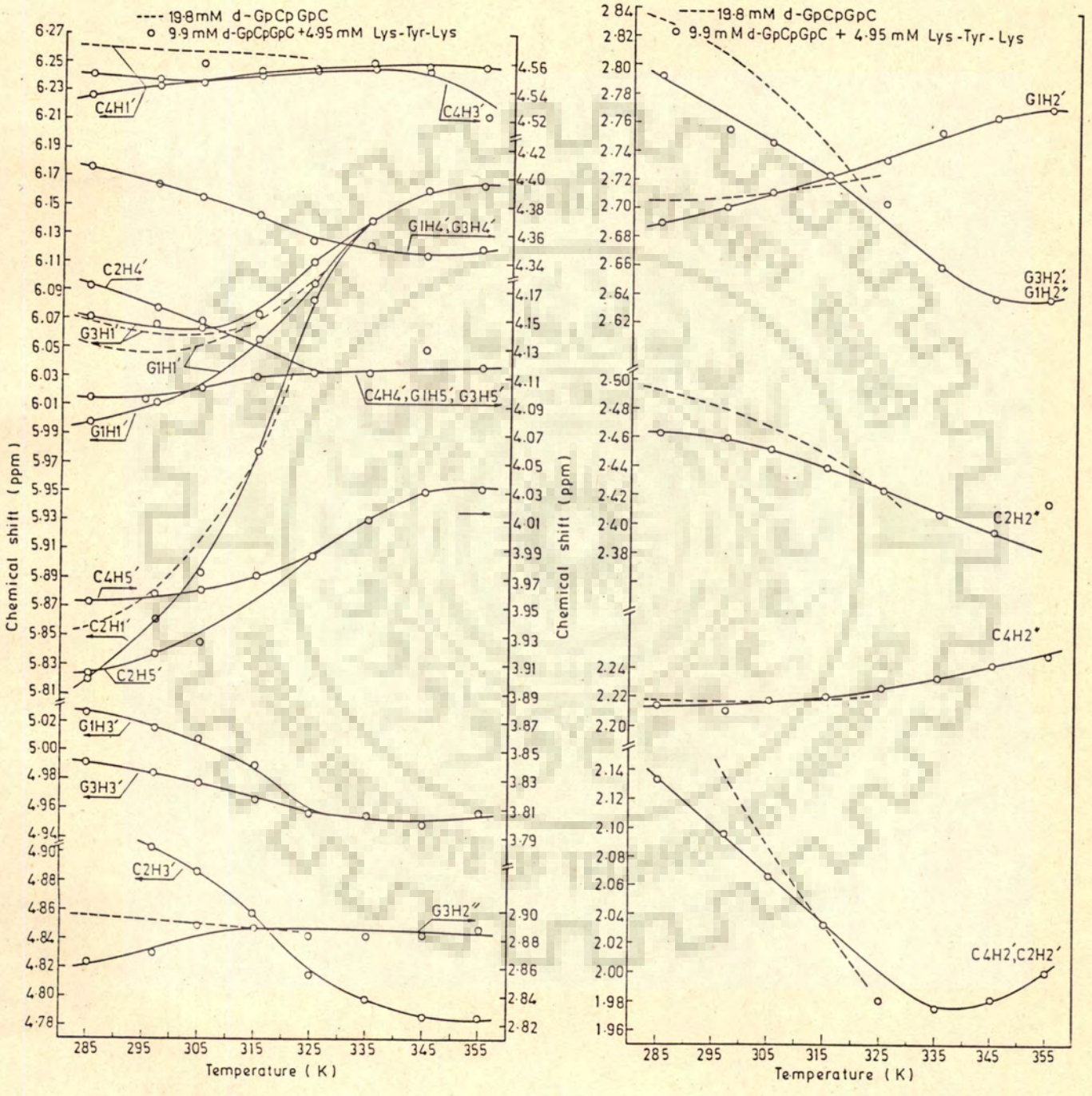

---- 19·8 mM d-GpCpGpC
o 9·9 mM d-GpCpGpC + 4·95 mM Lys-Tyr-Lys
C4H1'
C4H3'
G1H4', G3H4'
C2H4'
G3H1'
G1H1'
C4H4', G1H5', G3H5'
G1H1'
C4H5'
C2H1'
C2H5'
G1H3'
G3H3'
C2H3'
G3H2''
Chemical shift (ppm)
Temperature (K)
G1H2'
G3H2', G1H2*
C2H2*
C4H2*
C4H2', C2H2'

Fig. 4.8 Contd.

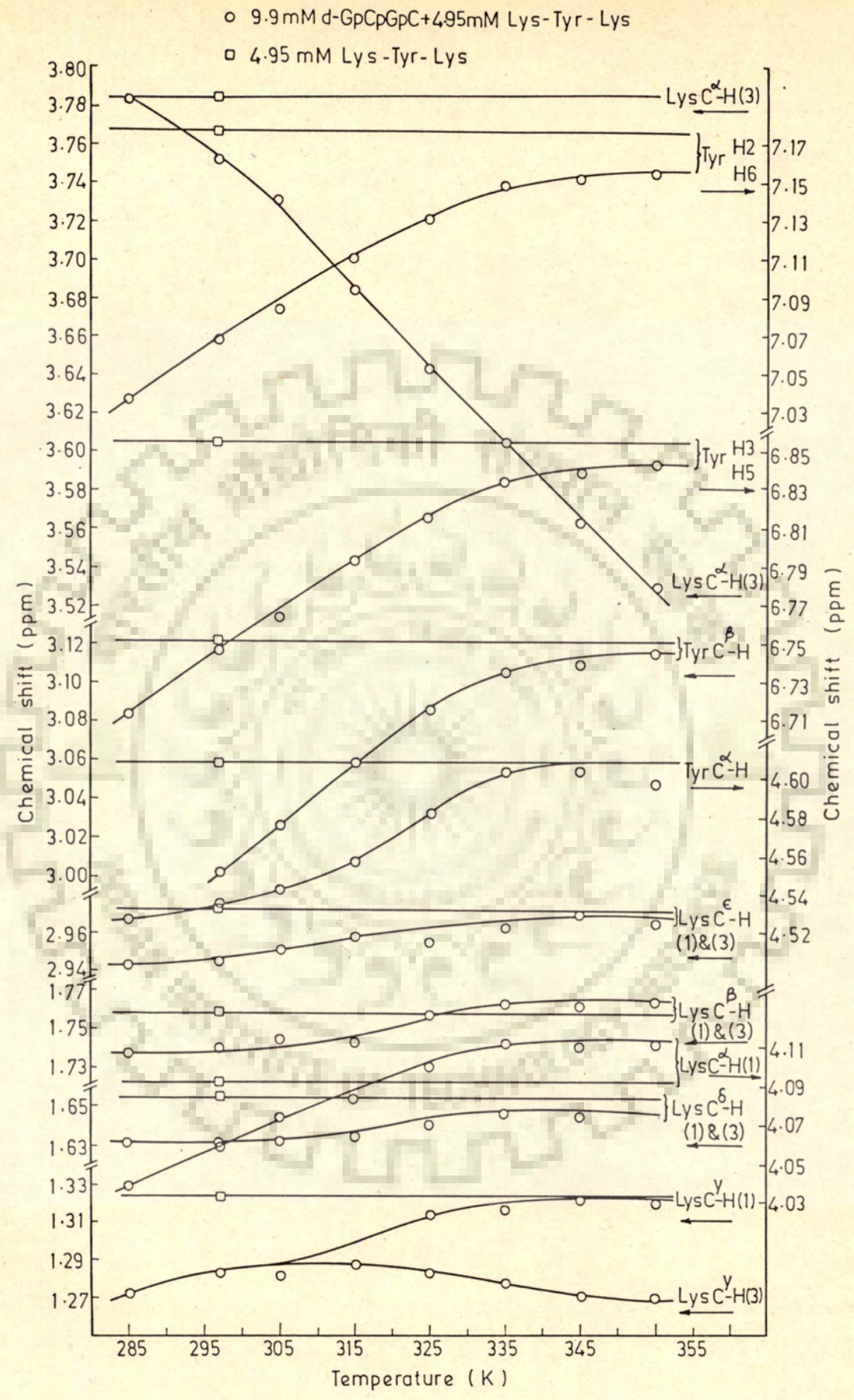

o 9.9 mM d-GpCpGpC+4.95mM Lys-Tyr-Lys
□ 4.95 mM Lys-Tyr-Lys
Lys Cα-H(3)
Tyr H2
H6
Tyr H3
H5
Lys Cα-H(3)
Tyr Cβ-H
Tyr Cα-H
Lys Cε-H (1)&(3)
Lys Cβ-H (1)&(3)
Lys Cα-H(1)
Lys Cδ-H (1)&(3)
Lys Cγ-H(1)
Lys Cγ-H(3)
Chemical shift (ppm)
Chemical shift (ppm)
Temperature (K)

Fig. 4.8 Contd.

Table 4.2 : Changes in chemical shift, $\Delta\delta$ in ppm, of peptide and nucleotide protons on binding of 9.9 mM d-GpCpGpC to 4.95 mM Lys-Tyr-Lys at indicated temperatures

SYSTEM	Temp. (K)	G1H8	G3H8	C2H6	C4H6	C2H5	C4H5	G1H1'	G3H1'	C2H1'
d-GpCpGpC	285	8.235	8.001	7.475	7.522	5.510	5.572	6.054	6.067	5.857
	297	8.210	7.983	7.470	7.527	5.525	5.594	6.046	6.060	5.885
	315	8.160	7.958	7.491	7.595	5.593	5.682	6.069	6.069	5.976
d-GpCpGpC + Lys-Tyr-Lys	285	8.232	7.977	7.469	7.489	5.475	5.521	5.999	6.071	5.819
	297	8.183	7.963	7.505	7.459	5.505	5.561	6.012	6.064	5.861
	315	8.120	7.950	7.496	7.601	5.615	5.695	6.057	6.073	5.980
$\Delta\delta$ (ppm)*	285	0.003	0.024	0.006	0.033	0.035	0.051	0.055	-0.004	0.038
	297	0.027	0.020	-0.035	0.068	0.020	0.033	0.034	-0.004	0.024
	315	0.040	0.008	-0.005	-0.006	-0.022	-0.013	0.012	-0.004	-0.004

SYSTEM	Temp. (K)	C4H1'	$\frac{C2H2'}{C4H2'}$	C2H2"	C4H2"	G1H2"	$\frac{G3H2'}{G1H2"}$	G3H2"	G1H3'	G3H3'
d-GpCpGpC	285	6.260	2.219	2.493	2.219	2.708	2.831	2.895	5.024	5.024
	297	6.259	2.128	2.478	2.225	2.710	2.806	2.893	5.014	5.014
	315	6.255	2.031	2.441	2.218	2.737	2.737	2.888	-	-
d-GpCpGpC + Lys-Tyr-Lys	285	6.226	2.133	2.462	2.214	2.689	2.792	2.863	5.026	4.991
	297	6.232	2.096	2.460	2.210	2.684	2.755	2.870	5.016	4.984
	315	6.242	2.032	2.438	2.219	2.721	2.721	2.887	-	-
$\Delta\delta$ (ppm)*	285	0.034	0.086	0.031	0.005	0.019	0.039	0.032	-0.002	0.033
	297	0.027	0.032	0.018	0.015	0.026	0.051	0.023	0.022	0.030
	315	0.013	-0.001	0.003	-0.001	0.016	0.016	0.001	-	-

Contd.

SYSTEM	Temp. (K)	C4H3'	G1H4'/ G3H4'	C2H4'	C4H4',G1H5'/ G3H5'	C2H5'	C4H5'	Tyr H2,H6	Tyr H3,H5
d-GpCpGpC / Lys-Tyr-Lys	285	4.549	4.424	4.170	4.113	3.858	3.948	-	-
	297	4.538	4.410	4.163	4.112	3.856	3.951	7.177	6.854
	315	4.532	4.370	4.115	4.115	3.893	3.934	-	-
d-GpCpGpC + Lys-Tyr-Lys	285	4.550	4.406	4.172	4.095	3.902	3.953	7.037	6.713
	297	4.547	4.394	4.157	4.094	3.917	3.957	7.069	6.748
	315	4.551	4.373	4.101	4.110	3.971	3.971	7.111	6.793
Δδ (ppm)*	285	-0.001	0.018	-0.002	0.018	-0.044	-0.005	0.140	0.141
	297	0.009	0.016	0.006	0.017	-0.061	-0.006	0.108	0.106
	315	-0.019	-0.003	0.014	0.005	-0.078	-0.037	0.066	0.061

SYSTEM	Temp. (K)	$Tyr C^{\alpha}$-H	$Tyr C^{\beta}$-H	$Lys C^{\alpha}$-H (1)	$Lys C^{\alpha}$-H (3)	$Lys C^{\beta}$-H (1)&(3)	$Lys C^{\gamma}$-H (1)&(3)	$Lys C^{\delta}$-H (1)&(3)	$Lys C^{\varepsilon}$-H (1)&(3)
Lys-Tyr-Lys	297	4.608	3.121	4.089	3.785	1.758	1.323	1.655	2.971
d-GpCpGpC + Lys-Tyr-Lys	285	4.570	-	4.040	3.783	1.738	1.272	1.632	2.944
	297	4.533	3.008	4.060	3.753	1.740	1.284	1.632	2.945
	315	4.558	3.059	4.086	3.685	1.745	1.288	1.635	2.958
Δδ(ppm)*	285	0.038	-	0.049	0.002	0.020	0.051	0.023	0.027
	297	0.075	0.113	0.029	0.032	0.018	0.039	0.023	0.026
	315	0.050	0.062	0.003	0.100	0.013	0.035	0.020	0.013

*Upfield shifts, Δδ in ppm, are taken with positive sign.

chemical shift between highest and lowest temperature has increased. The $T_{\frac{1}{2}}$ observed for bound complex varies between 319-324 K for various resonances as follows:

	$T_{\frac{1}{2}}$
C2H6	324 K
C4H6	319 K
G3H1'	323 K
C2H5	321 K
C4H5	320 K

The base protons GH8, CH5, CH6 all have practically shifted upfield upto ~0.05 ppm at 285 K on binding. However, the Tyr ring protons shift upfield significantly by about ~0.14 ppm at 285 K. Further, the change in chemical shift $\Delta\delta$ on binding decreases with temperature, it being largest at 285 K. It clearly demonstrates that binding preferentially occurs to the double-stranded d-GpCpGpC than to the single-strand. The large upfield shifts in Tyr ring protons is obviously due to stacking of Tyr ring between base pairs of d-GpCpGpC. It may be noted from the change in chemical shift of Tyr ring protons on binding to deoxyoligonucleotide that the binding to d-GpCpGpC is to a larger extent ($\Delta\delta$ ~ 0.14 ppm at 285 K) than that to d-CpG ($\Delta\delta$ ~ .014 ppm) at 285 K). Comparatively large value of $\Delta\delta$ that is, 0.02-0.05 ppm of lysine protons is likely due to the electrostatic interaction of its side chain with negatively charged DNA.

Figs. 4.9 and 4.10 show results of 2D-COSY and NOESY spectra of a mixture of d-GpCpGpC with Lys-Tyr-Lys for the same sample for which 1D-NMR results have been shown in Fig. 4.7-4.8. Figs. 4.9a and 4.10a show portions of expanded 2D-spectra to highlight specific connectivities. Figs. 4.11 and 4.12 are schematic representation of these results. Assignments of various peaks are easily made by using the uncomplexed spectra of d-GpCpGpC and Lys-Tyr-Lys as reference. For C4, G1 and G3 bases H1' shows J-connectivity both the corresponding H2' and H2"; H2' with corresponding H3'; H2" with corresponding H3'; and H3' with corresponding H4'. This clearly shows that C4, G1 and G3 exist in O1'-endo sugar conformation. For C2 base, H1' with H2' cross peak exists but H1' with H2" cross peak is not seen, however, the presence of H1' to H2' and H3' to H4' cross peaks conclude that the C2 exist in O1'-endo sugar conformation.

The relative intensities of NOE cross peaks in d-GpCpGpC bound to Lys-Tyr-Lys are as follows:

G1H8... G1H2'	intra-residue	strong	(31)
G1H8...G1H2"	-do-	weak	(32)
G1H8...G1H1'	-do-	weak	(33)
C2H6...G1H1'	inter-residue	Not seen	(34)
C2H6...G1H2"	-do-	weak	(35)

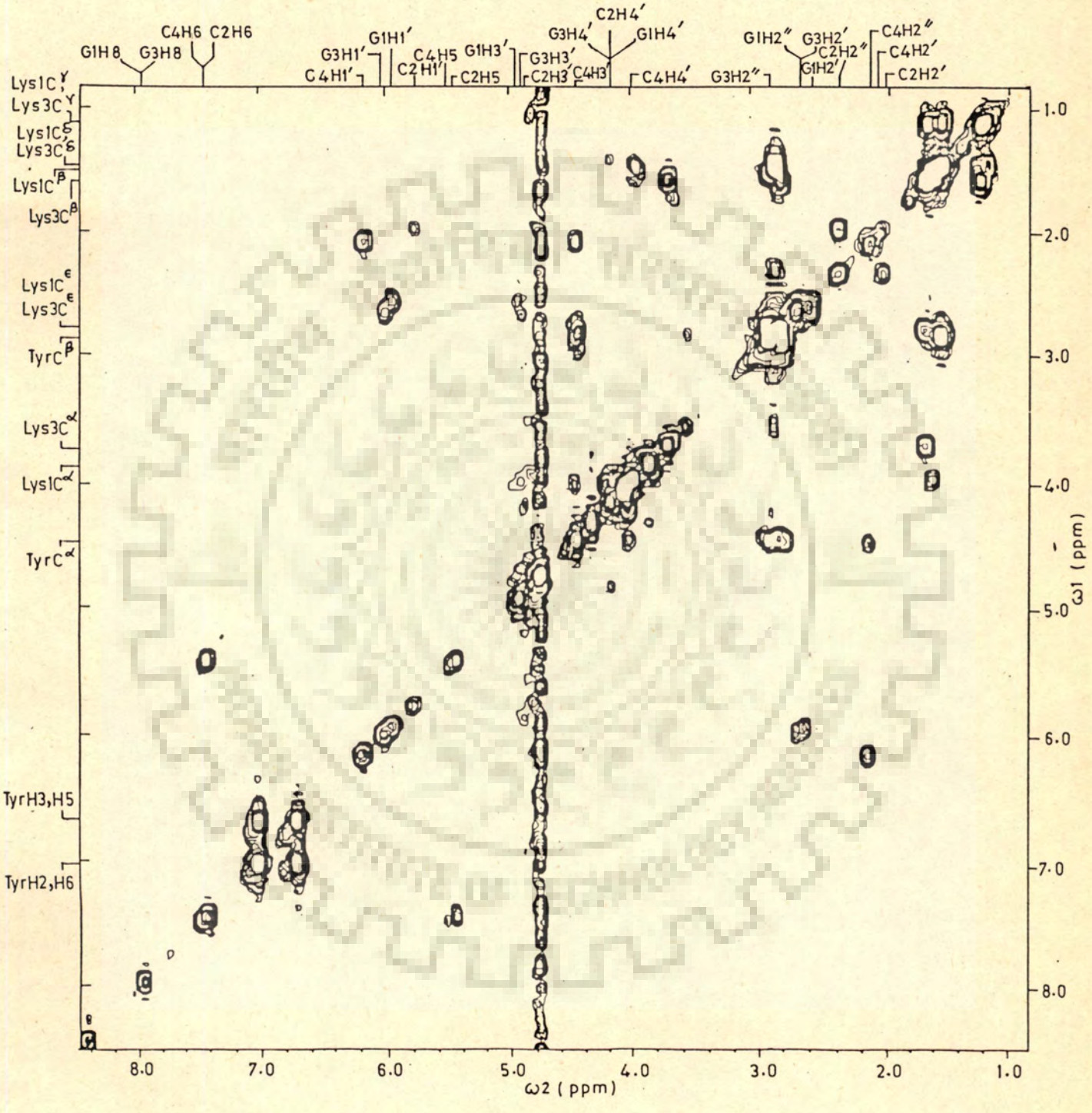


Fig. 4.9 500 MHz COSY spectrum of a mixture 9.9 mM d-GpCpGpC and 4.95 mM Lys-Tyr-Lys in D_2O at 297 K for the sample used in Fig. 4.7.

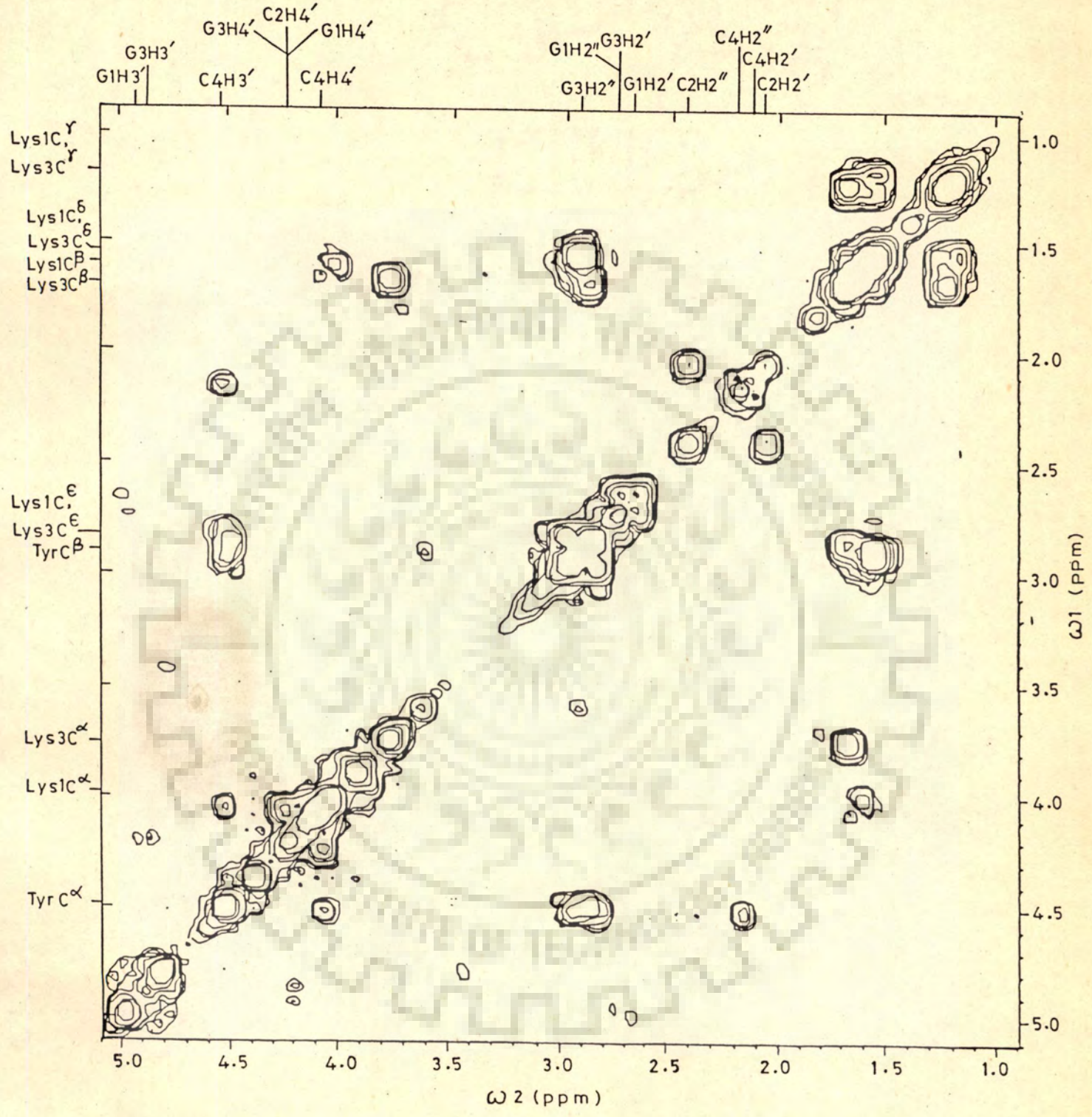


Fig. 4.9 (a) Portion of COSY spectrum of Fig. 4.9 in an expanded scale.

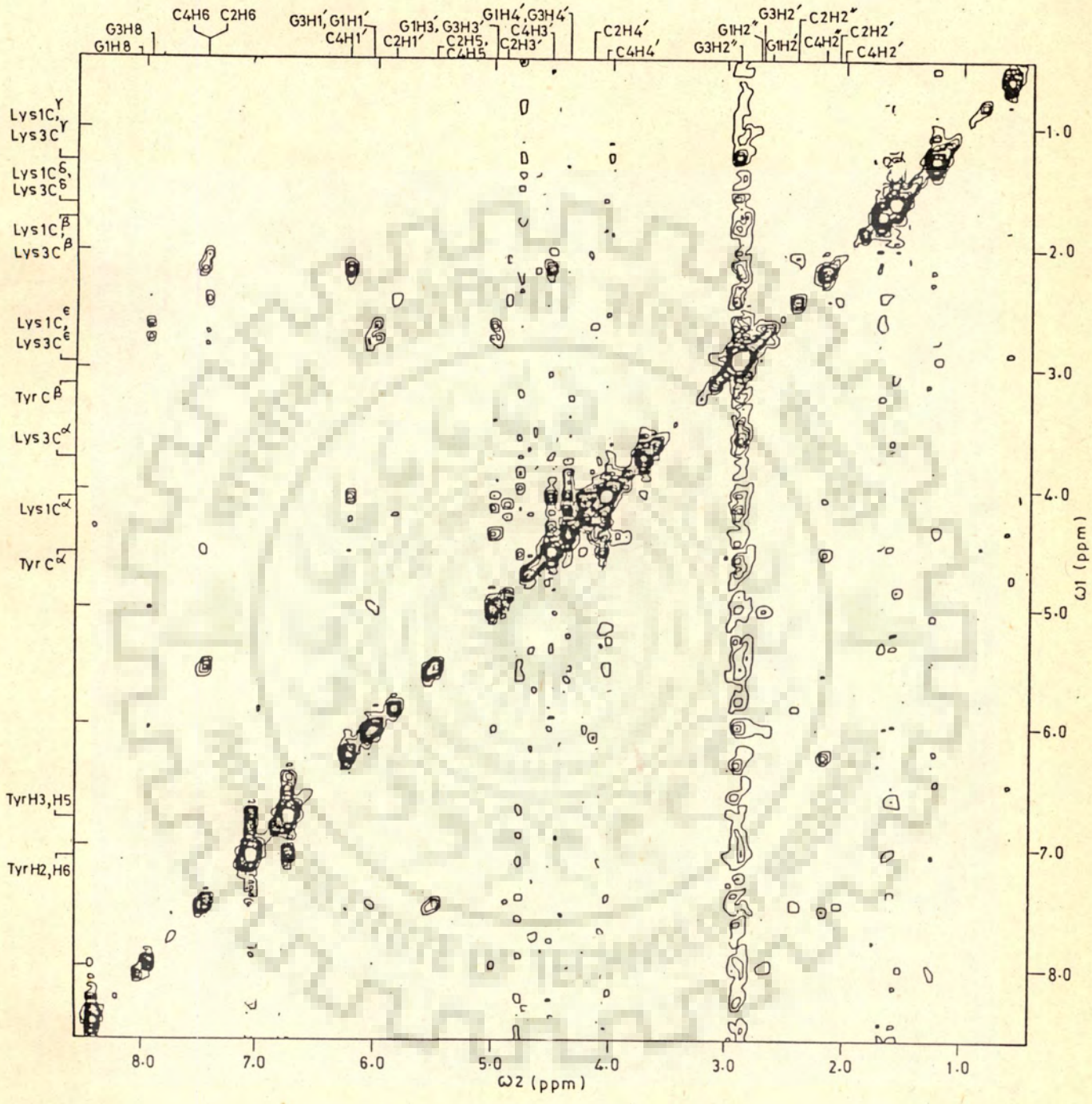


Fig. 4.10 500 MHz NOESY spectrum of a mixture of 9.9 mM d-GpCpGpC and 4.95 mM Lys-Tyr-Lys in D_2O at 297 K for the sample used in Fig. 4.7.

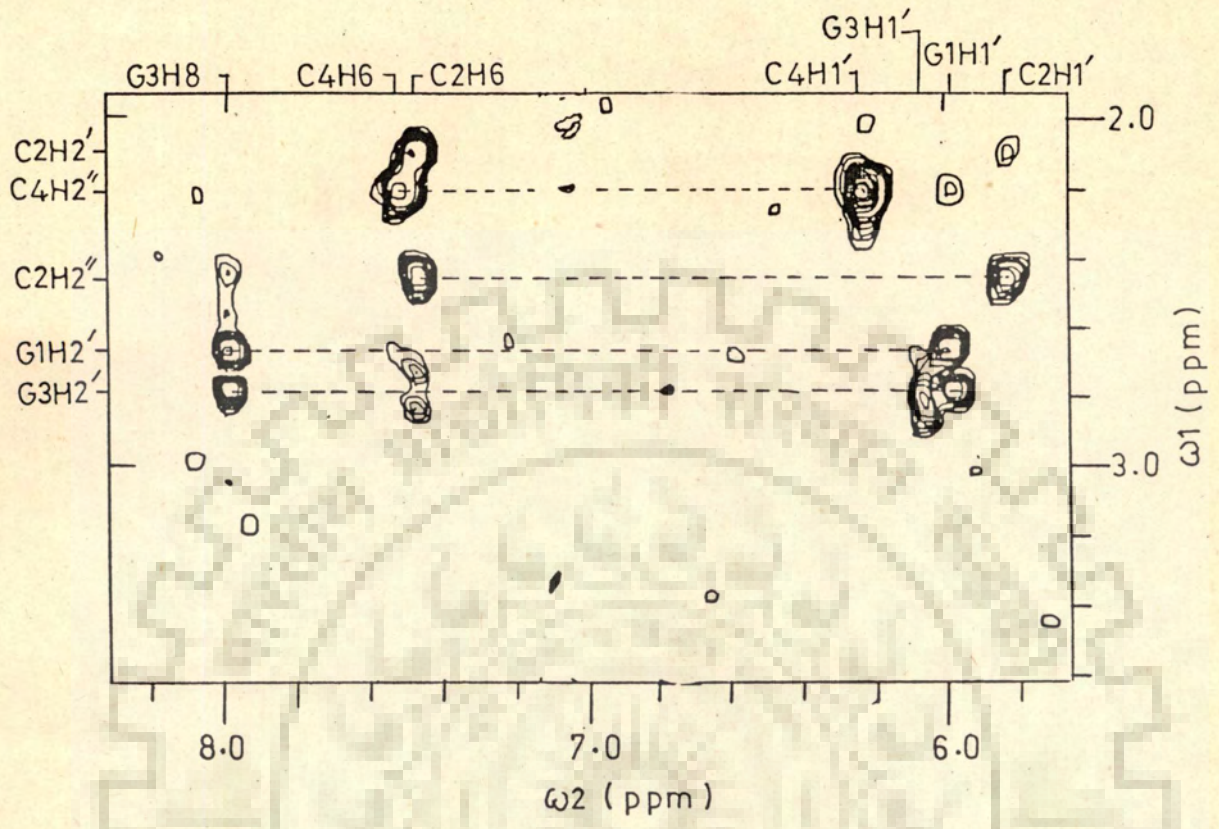


Fig. 4.10 (a) Portion of NOESY spectrum of Fig.4.10 in an expanded scale

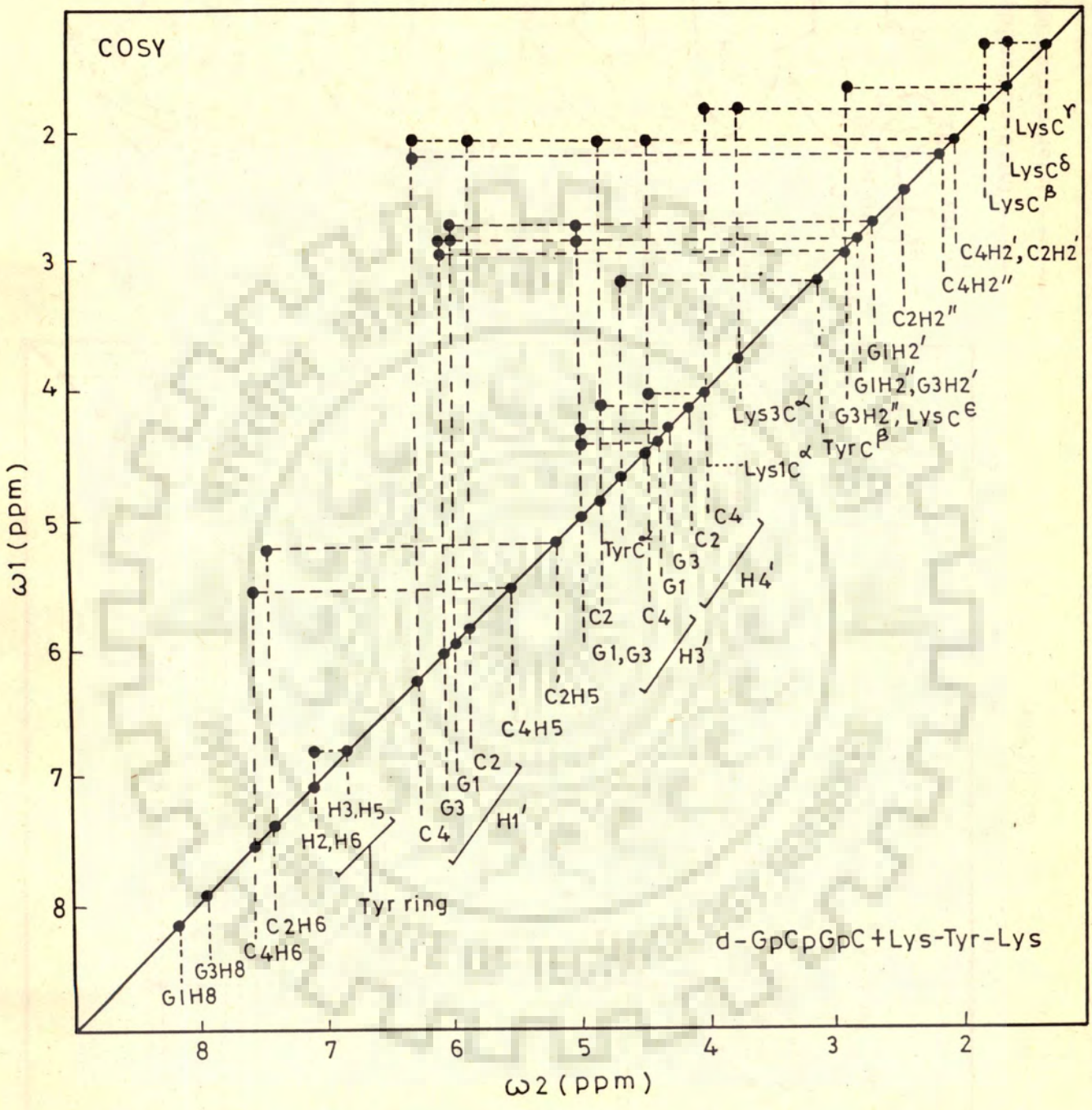


Fig. 4.11 Schematic representation of the results of COSY spectra of mixture of d-GpCpGpC and Lys-Tyr-Lys taken from Fig.4.9

C2H6... C2H2'	intra-residue	strong	(36)
C2H6... C2H2"	-do-	strong	(37)
C2H6... C2H1'	-do-	Not seen	(38)
G3H8... C2H1'	inter-residue	Not seen	(39)
G3H8... C2H2"	-do-	strong	(40)
G3H8... G3H2'	intra-residue	strong	(41)
G3H8... G3H2"	-do-	weak	(42)
G3H8... G3H1'	-do-	weak	(43)
C4H6... G3H1'	inter-residue	strong	(44)
C4H6... G3H2"	-do-	weak	(45)
C4H6... C4H2'	intra-residue	strong	(46)
C4H6... C4H2"	-do-	strong	(47)
C4H6... C4H1'	-do-	Not seen	(48)
G1H1'...G1H2'	-do-	weak	(49)
G1H1'...G1H2"	-do	strong	(50)
C2H1'...C2H2'	-do-	not seen	(51)
C2H1'...C2H2"	-do-	strong	(52)
G3H1'...G3H2'	-do-	weak	(53)
G3H1'...G3H2"	-do-	strong	(54)
C4H1'...C4H2'	-do-	weak	(55)
C4H1'...C4H2"	-do-	strong	(56)

C4H6 ...G1H2"	inter-strand	weak	(57)
C4H2"...Tyr C^{α}	intermolecular	strong	(58)
C4H1'...Lys C^{α}	-do-	weak	(59)

For both G1 and G3, the NOE connectivity of base H8 proton to corresponding H2' is stronger [(31) and (41)] than that to its H2" [(32) and (42)] indicating that both G1 and G3 have χ_{CN} in anti conformation. For C2 and C4, the base to H2' and H2" NOE's are equally strong and they could be in anti or high anti χ_{CN} conformation. A weak inter-strand NOE between C4H6 and G1H2" (No. 57) indicates that tetranucleotide exists as double-helix. Strong intra-residue NOE's of base protons to their corresponding H2' protons [No.(31), (36), (41) and (46)] and inter-residue (NOE'S No. (40), (45), (35) indicate that the DNA exists a right-handed helix. The appearance of intermolecular NOE's between tyrosine C^{α}, Lys C^{α}, and sugar protons [(No.(58) and (59))] indicate proximity of peptide to the tetranucleotide in the complex.

The conformation of the complex of d-GpCpGpC with Lys-Tyr-Lys therefore is, C2, C4, G1, G3 sugar with O1'-endo pucker; G1 and G3 bases with χ_{CN} as anti; C2 and C4 bases with χ_{CN} as anti/high anti; and the nucleic acid helix sense as right-handed.

CHAPTER - 5

d-GpC AND ITS BINDING TO Lys-Trp-Gly-Lys OtBu AND Lys-Tyr-Lys

500 MHz proton NMR spectra of deoxyoligonucleotide d-GpC and its complex with tetrapeptide Lys-Trp-Gly-Lys OtBu and tripeptide Lys-Tyr-Lys as a function of temperature in the range 277-345 K has been studied. Homonuclear 2D-COSY and NOESY spectra of the samples were taken at 297 K to find their conformation.

d-GpC

Figures 5.1 and 5.2 show the results of NMR spectra of d-GpC in D_2O as a function of temperature. The chemical shift values are listed in Table 5.1. The assignments are based on the known positions occupied by these protons in oligonucleotides and on the basis of 2D-COSY and NOESY spectra given in Figures 5.3 and 5.4. The $T_{\frac{1}{2}}$ value for the helix to coil transition on increasing temperature is obtained as 286 K for CH5 and CH6 protons. A schematic representation shown in Figure 5.5 summarises the results of 2D-NMR experiments. Since the J—connectivities between H1'-H2', H1'-H2", H2'-H3', H3'-H4' were found to exist for both the nucleotide residues the sugar conformations are 01'-endo for both G and C. The NOESY spectra is not well

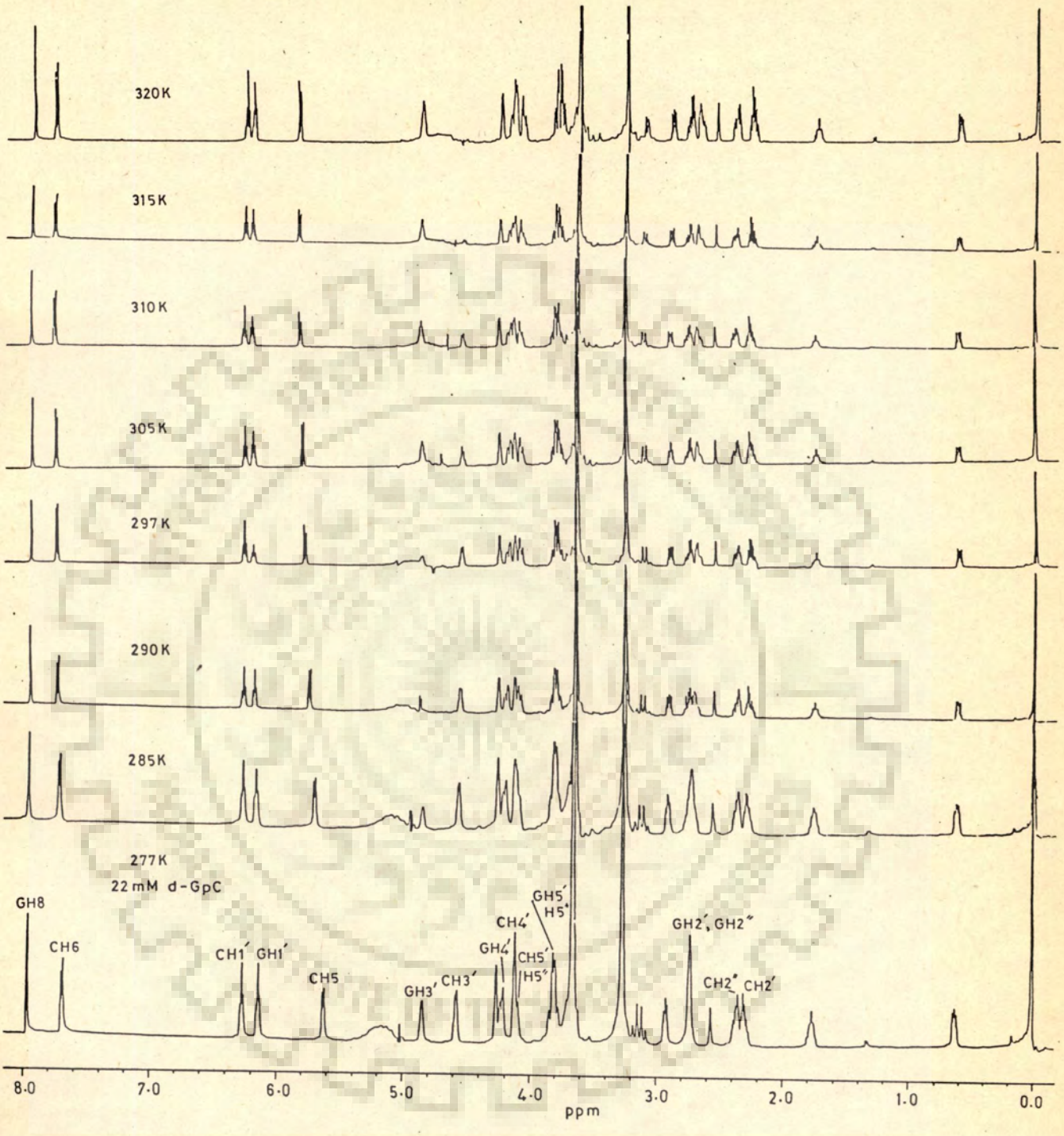


Fig. 5.1 500 MHz proton NMR spectra of 22 mM (strand concentration) deoxydinucleotide d-GpC solution in D_2O (pH=7.0) containing 0.25 mM EDTA, at various indicated temperatures. Ref. DSS.

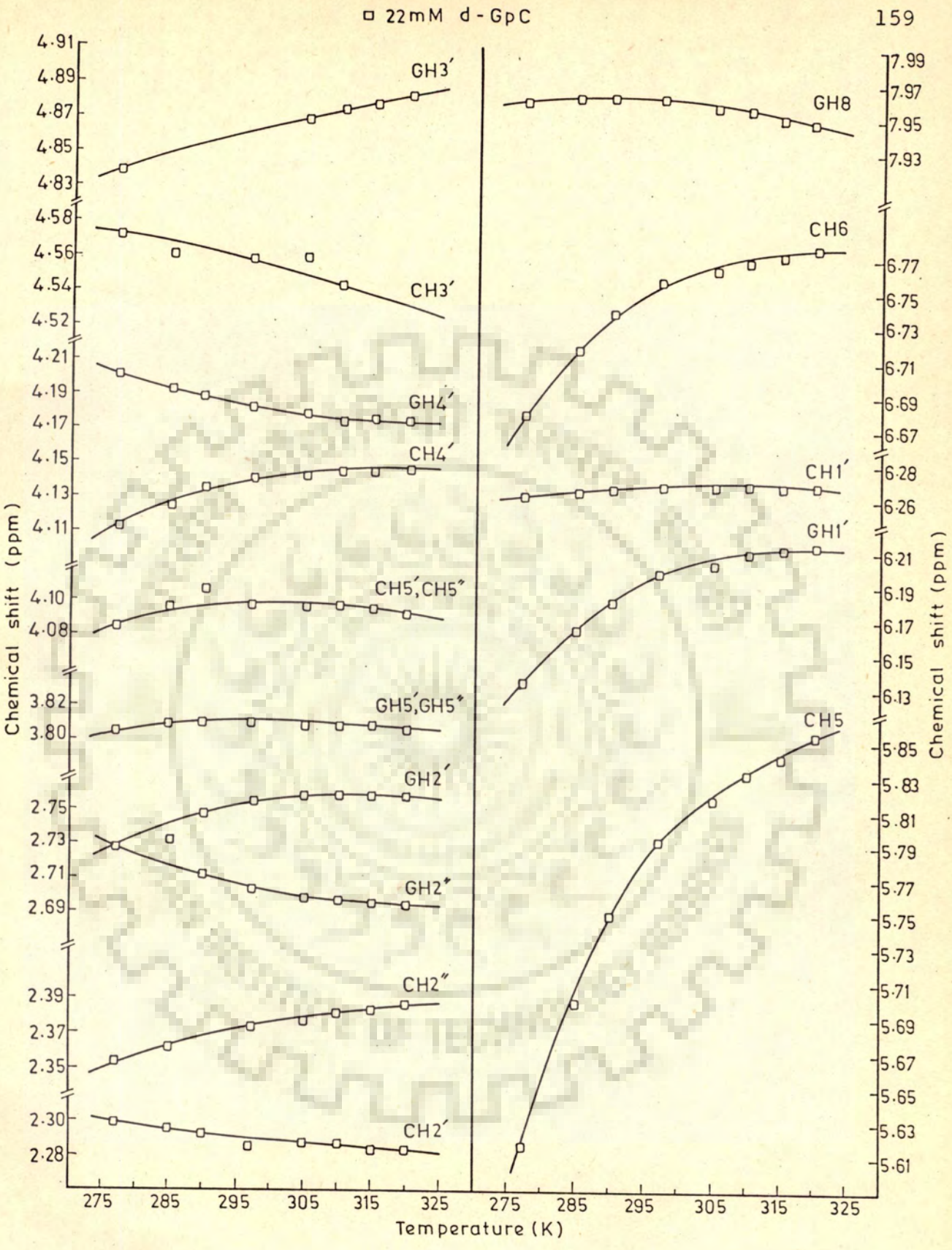


Fig. 5.2 Chemical shifts of base and sugar protons of 22 mM d-GpC as a function of temperature from spectra in fig. 5.1

Table 5.1 : Chemical shift values of various protons at 500 MHz for 22 mM d-GpC in D_2O (pH=7.0) in temperature range 277K-320K.

Temp. (K)	GH8	CH6	CH5	CH1'	GH1'	CH2'	CH2"	GH2'	GH2"
277	7.960	7.680	5.617	6.261	6.134	2.297	2.353	2.727	2.727
285	7.962	7.717	5.699	6.264	6.161	2.294	2.361	2.732	2.732
290	7.962	7.738	5.748	6.266	6.180	2.292	2.367	2.748	2.712
297	7.960	7.756	5.793	6.268	6.197	2.284	2.373	2.754	2.703
305	7.956	7.763	5.813	6.267	6.202	2.287	2.377	2.758	2.698
310	7.954	7.769	5.830	6.268	6.208	2.287	2.372	2.759	2.696
315	7.950	7.772	5.842	6.266	6.210	2.284	2.383	2.758	2.694
320	7.949	7.776	5.854	6.267	6.213	2.284	2.387	2.758	2.694

Temp. (K)	CH3'	GH3'	CH4'	GH4'	GH5'/H5"	CH5'/H5"
277	4.571	4.837	4.113	4.201	3.804	4.084
285	4.560	-	4.124	4.193	3.808	4.096
290	4.562	-	4.135	4.189	3.809	4.107
297	4.558	4.869	4.141	4.181	3.810	4.097
305	4.559	4.871	4.141	4.178	3.808	4.096
310	4.542	4.874	4.144	4.172	3.807	4.096
315	-	4.876	4.144	4.176	3.809	4.093
320	-	4.881	4.146	4.174	3.805	4.093

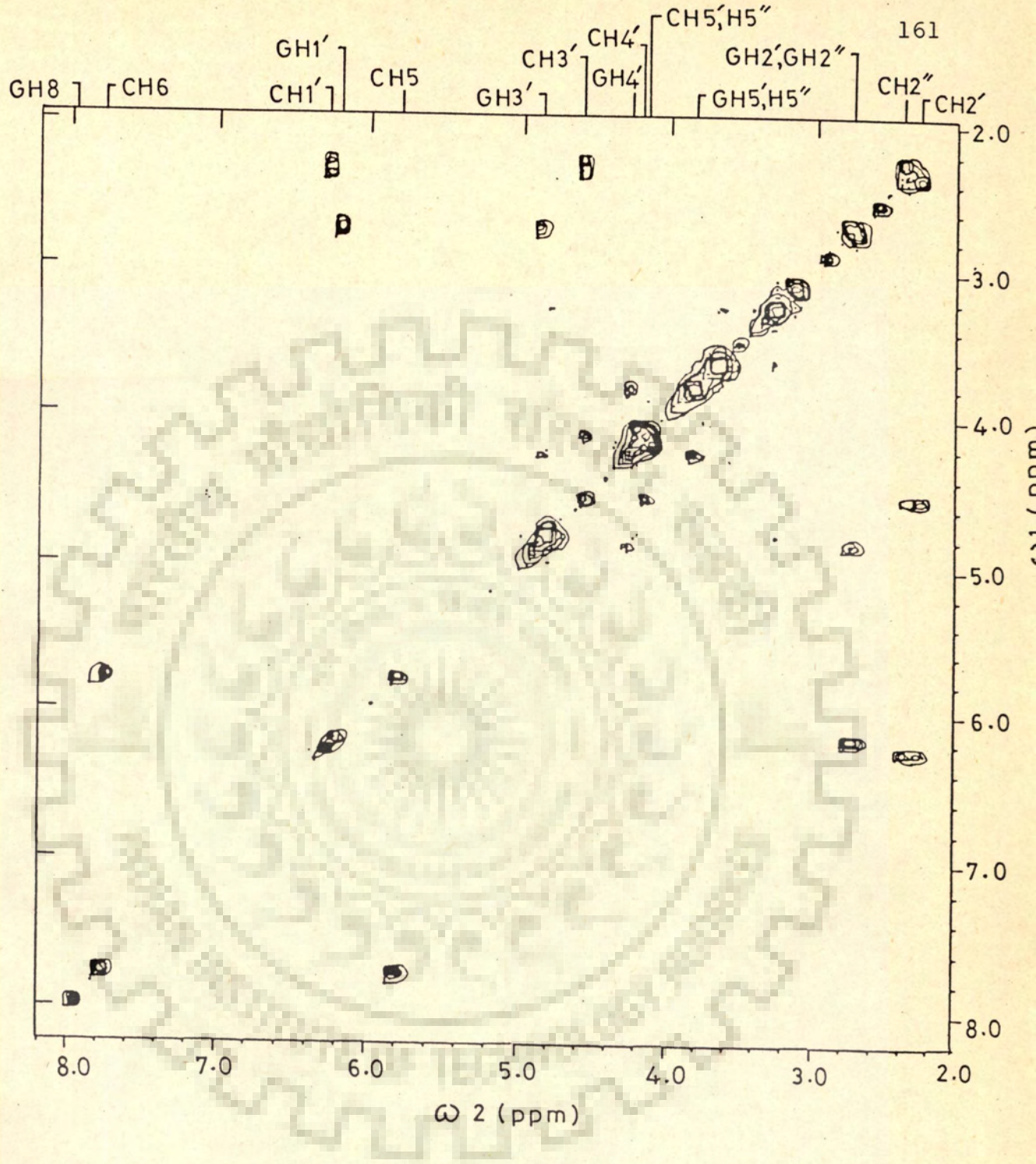


Fig. 5.3 500 MHz COSY spectrum of 22 mM d-GpC solution in D_2O containing 0.25 mM EDTA (pH=7.1) at 297 K.

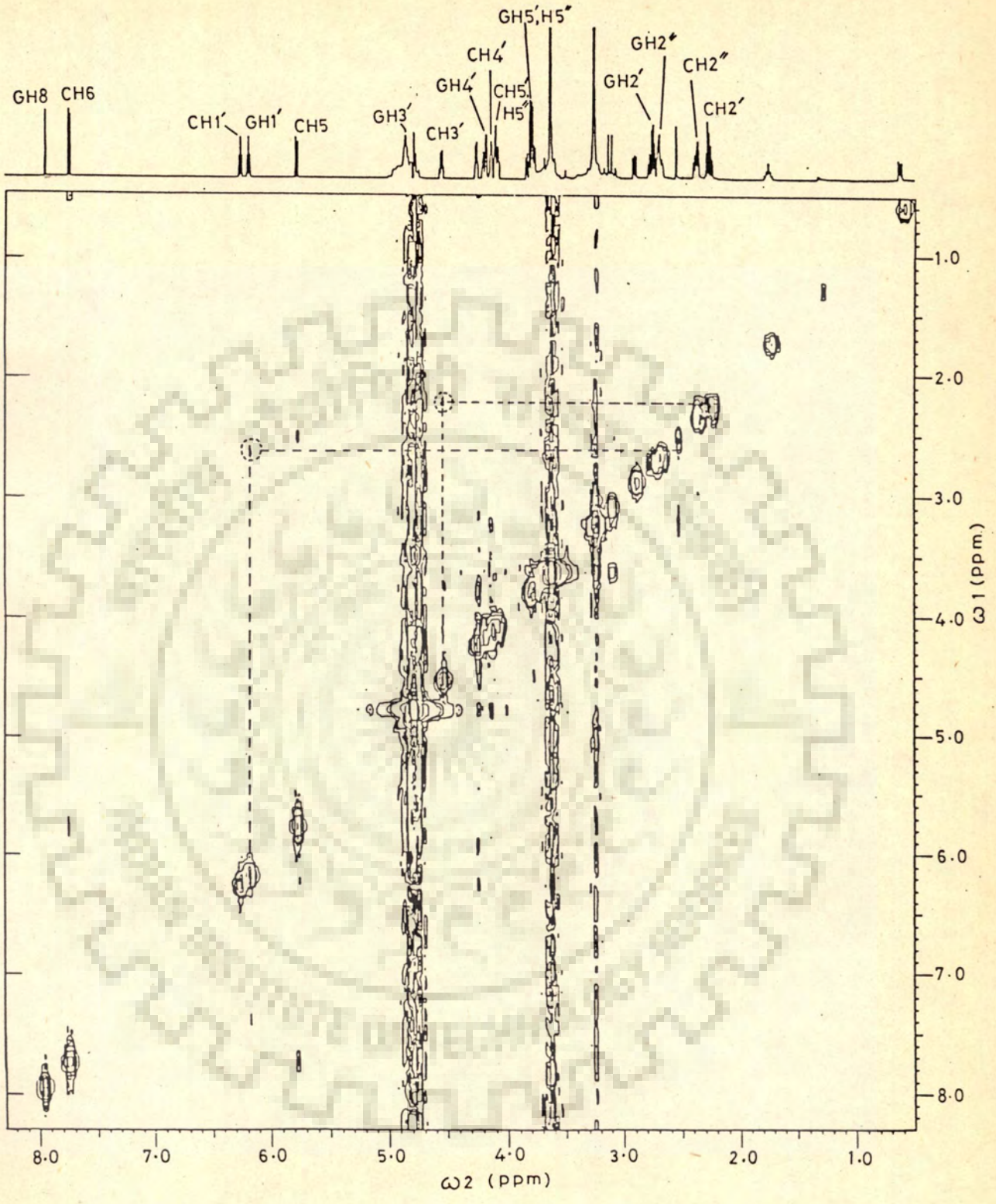


Fig. 5.4 500 MHz NOESY spectrum of 22 mM d-GpC solution in D_2O containing 0.25 mM EDTA (pH=7.1) at 297 K

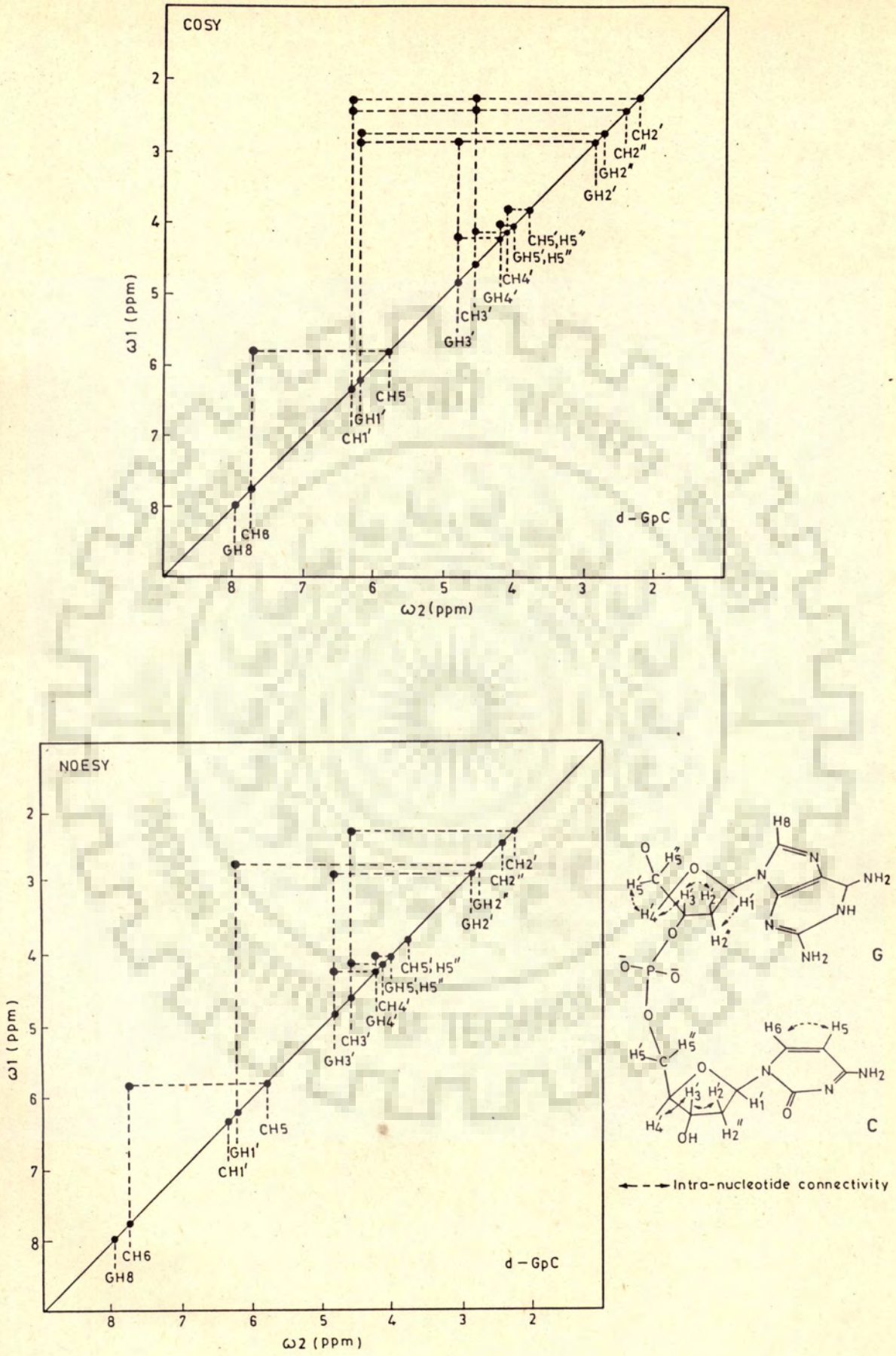


Fig. 5.5 Schematic representation of the results of COSY and NOESY spectra of d-GpC shown in Figs. 5.3 and 5.4.

resolved and the connectivities GH1'-GH2", GH2'-GH3', GH3'-GH4', CH3'-CH4', GH4'-GH5' seen in figure do not enable us to find more details about conformation of d-GpC.

BINDING OF d-GpC TO Lys-Trp-Gly-Lys OtBu

Figures 5.6, 5.7a, and 5.7b show the results of binding of d-GpC to tetrapeptide Lys-Trp-Gly-Lys OtBu at different temperatures. Changes in chemical shift of all protons are given in Table 5.2. It is found that $T_{\frac{1}{2}}$ value on binding increases from 286 K to 303-305 K. Trp ring protons shift upfield by about 0.05 ppm at 285 K. This change in chemical shift decreases with temperature showing that binding is preferential to double-helix. The upfield shift is due to stacking/intercalation between base pairs of d-GpC. It is comparable to that found on binding to d-CpG.

Figures 5.8 and 5.9 show results of COSY and NOESY spectra of a mixture of d-GpC with Lys-Trp-Gly-Lys OtBu for the same sample solution for which 1D NMR have been shown. Figures 5 8a and 5.9a are the portions of these spectra on expanded scales and Figure 5.10 is schematic representation of these results. For both the bases, the connectivities with various sugar resonances are found to be the same as that in uncomplexed state. The sugar conformation is, therefore, predominantly O1'-endo.

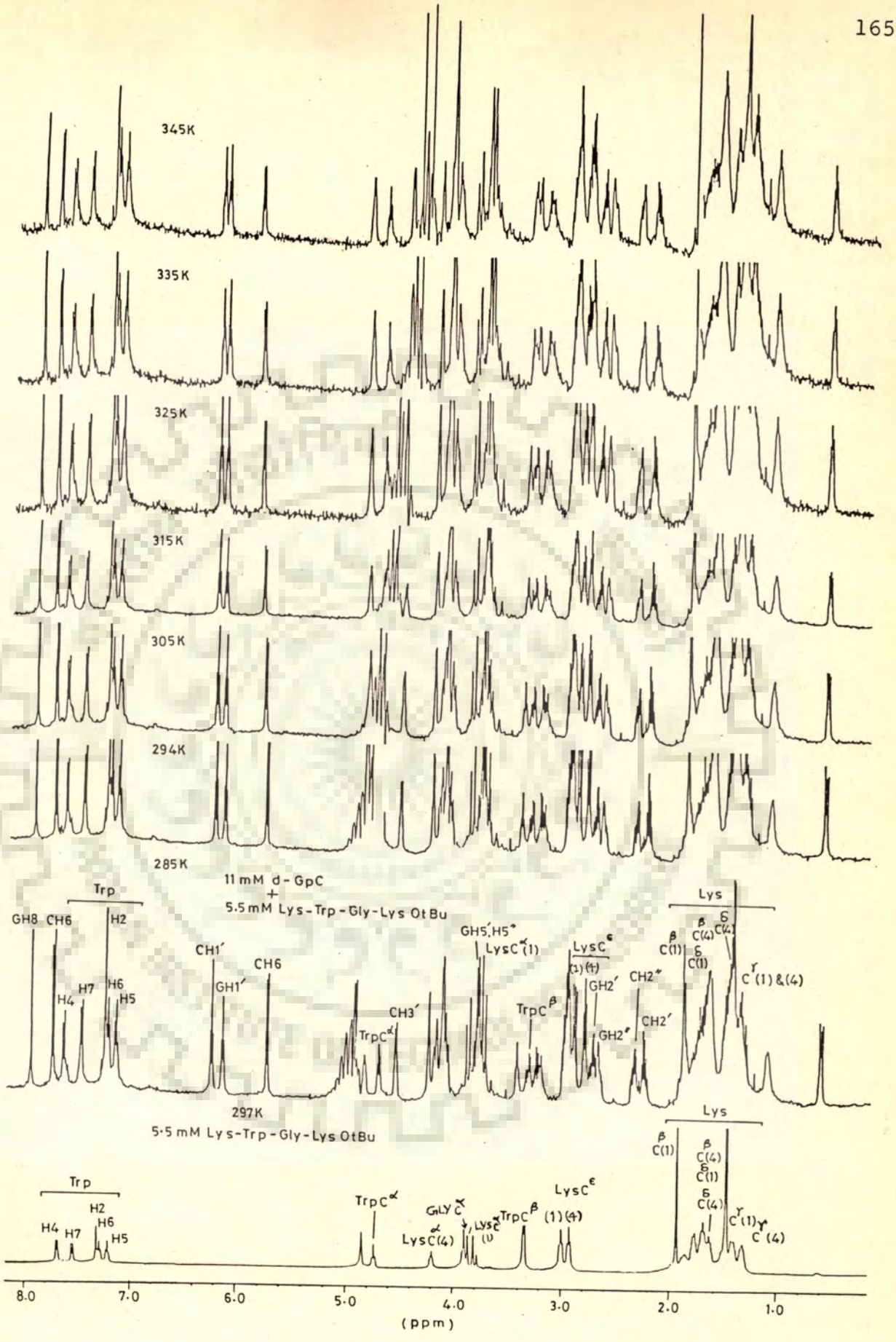


Fig. 5.6 500 MHz proton NMR spectra of a mixture of 11 mM d-GpC and 5.5 mM Lys-Trp-Gly-Lys OtBu at different temperatures in D_2O (pH=7.0) containing 0.25 mM EDTA. For comparison 500 MHz NMR spectrum of 5.5 mM Lys-Trp-Gly-Lys OtBu in D_2O solution (pH=7.0) at 297K is also shown. Ref. DSS.

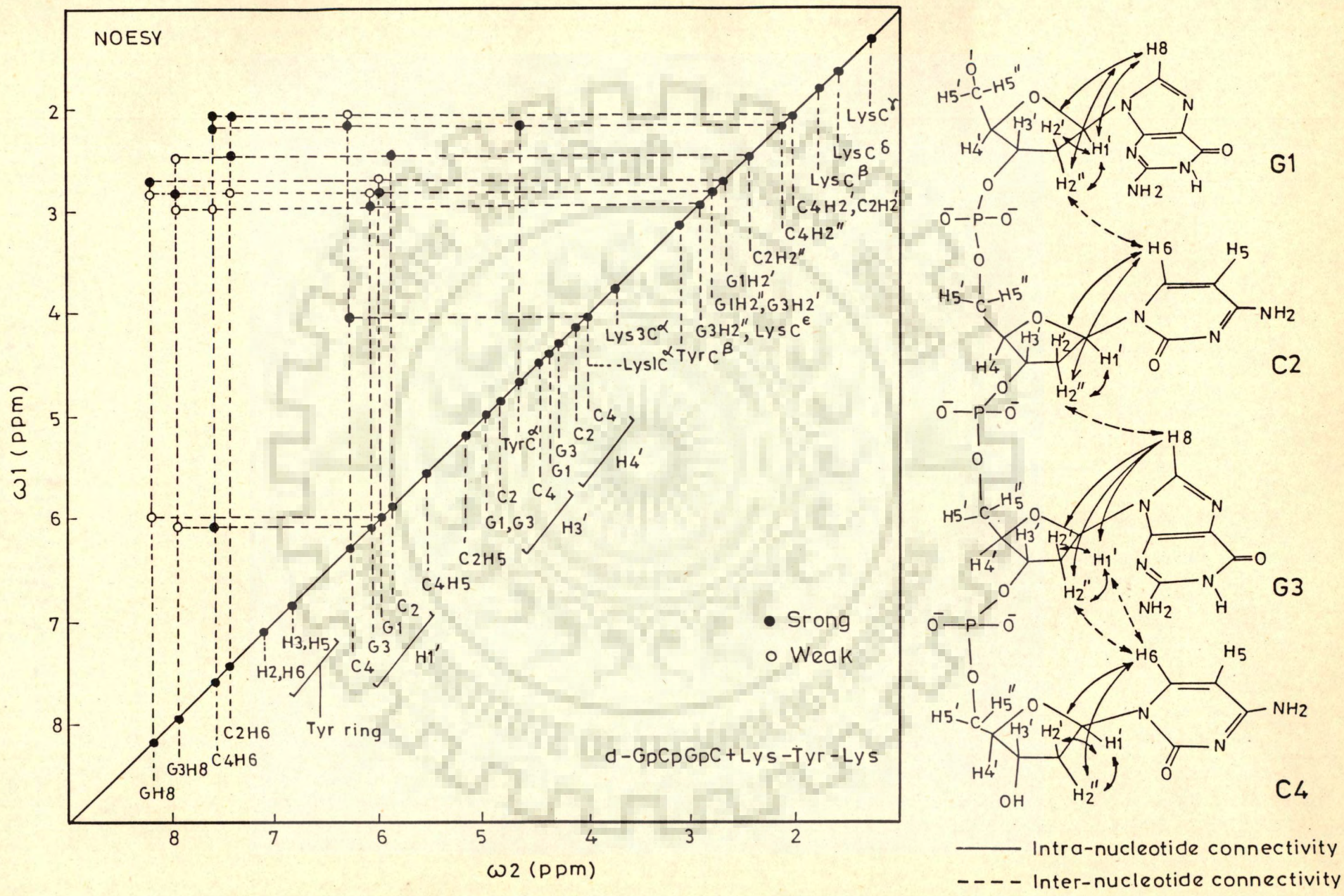


Fig. 4.12 Schematic representation of the results of NOESY spectra of mixture of d-GpCpGpC and Lys-Tyr-Lys taken from Fig.4.10.

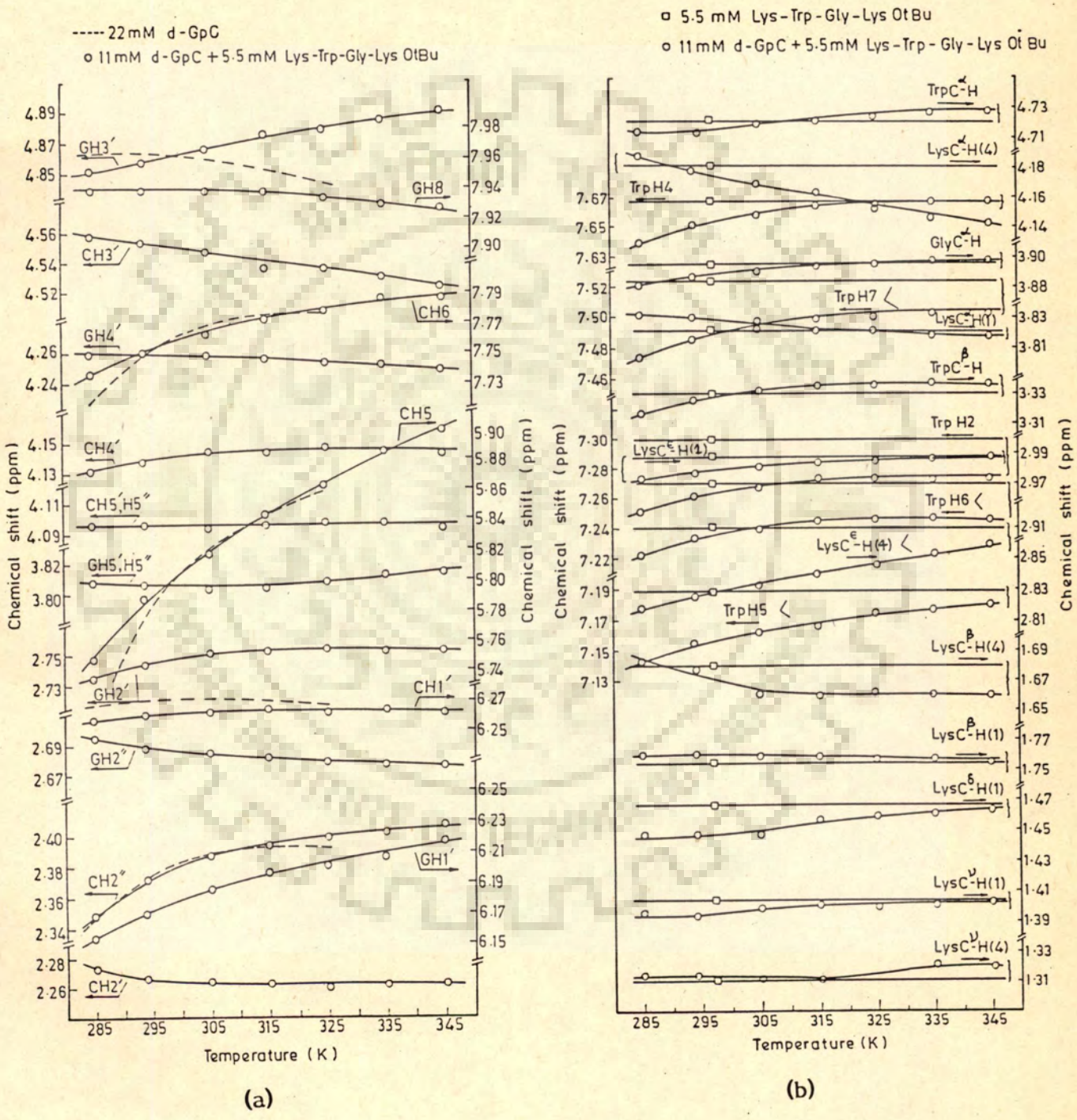


Fig. 5.7 (a), (b) Chemical shifts of some of the nucleotide and peptide protons as a function of temperature in uncomplexed nucleotide/peptide and in complexed state taken from data in Fig. 5.6.

Table 5.2 : Changes in chemical shift, $\Delta\delta$ in ppm, of peptide and nucleotide protons on binding of 11 mM d-GpC to 5.5 mM Lys-Trp-Gly-Lys OtBu at indicated temperatures

SYSTEM	Temp. (K)	GH8	CH6	CH5	CH1'	GH1'	CH2'	CH2"	GH2'	GH2"
d-GpC	285	7.962	7.717	5.699	6.264	6.161	2.294	2.361	2.732	2.732
	297	7.960	7.756	5.793	6.268	6.197	2.284	2.373	2.754	2.703
	315	7.950	7.772	5.842	6.266	6.210	2.284	2.383	2.758	2.694
d-GpC + Lys-Trp-Gly-Lys	285	7.940	7.736	5.747	6.257	6.154	2.273	2.348	2.733	2.695
	297	7.938	7.752	5.787	6.260	6.171	2.266	2.372	2.742	2.688
	315	7.938	7.772	5.843	6.265	6.196	2.263	2.394	2.753	2.685
$\Delta\delta$ (ppm)*	285	0.022	-0.019	-0.048	0.007	0.007	0.021	0.013	-0.001	0.037
	297	0.022	0.004	0.006	0.008	0.026	0.018	0.001	0.012	0.015
	315	0.012	0.000	-0.001	0.001	0.014	-0.021	-0.011	0.005	0.009

SYSTEM	Temp. (K)	CH3'	GH3'	CH4'	GH4'	CH5'/H5"	GH5'/H5"
d-GpC	285	4.560	-	4.124	4.193	4.096	3.808
	297	4.558	-	4.141	4.181	4.097	3.810
	315	-	4.876	4.144	4.176	4.093	3.809
d-GpC + Lys-Trp-Gly-Lys	285	4.558	-	4.132	4.259	4.096	3.808
	297	4.553	-	4.138	4.259	4.096	3.806
	315	-	4.874	4.143	4.255	4.100	3.805
$\Delta\delta$ (ppm)*	285	0.002	-	-0.008	-0.066	0.000	0.000
	297	0.005	-	0.003	-0.078	0.001	0.004
	315	-	0.002	0.001	-0.079	-0.007	0.004

Contd.

SYSTEM	Temp. (K)	Trp H4	Trp H7	Trp H2	Trp H5	Trp H6	Trp C^{α}-H	Trp C^{β}-H	Lys (1) C^{α}-H	Lys (4) C^{α}-H
Lys-Trp-Gly-Lys	297	7.668	7.524	7.300	7.190	7.270	4.721	3.330	3.779	4.181
d-GpC + Lys-Trp-Gly-Lys	285	7.639	7.474	7.251	7.145	7.222	4.714	3.317	3.830	4.188
	297	7.651	7.486	7.262	7.155	7.233	4.711	3.326	3.829	4.177
	315	7.664	7.498	7.273	7.167	7.245	4.726	3.335	3.821	4.162
Δδ (ppm)*	285	0.029	0.050	0.049	0.045	0.048	0.007	0.013	-0.051	-0.007
	297	0.017	0.038	0.038	0.035	0.037	0.010	0.004	-0.050	0.004
	315	0.004	0.026	0.027	0.023	0.025	0.001	-0.005	-0.042	0.019

SYSTEM	Temp. (K)	Lys (1) C^{β}-H	Lys (4) C^{β}-H	Lys (1) C^{γ}-H	Lys (4) C^{γ}-H	Lys (1)&(4) C^{δ}-H	Lys (1) C^{ε}-H	Lys (4) C^{ε}-H	Gly C^{α}-H
Lys-Trp-Gly-Lys	297	1.755	1.680	1.405	1.311	1.467	2.986	2.911	3.882
d-GpC + Lys-Trp-Gly-Lys	285	1.759	1.682	1.396	1.314	1.446	2.974	2.917	3.880
	297	1.759	1.677	1.393	1.313	1.447	2.977	2.909	3.886
	315	1.758	1.688	1.400	1.311	1.457	2.983	2.904	3.893
Δδ (ppm)*	285	-0.004	-0.002	0.009	-0.003	0.021	0.012	0.000	0.002
	297	-0.004	0.003	0.012	-0.002	0.020	0.009	0.002	-0.004
	315	-0.003	-0.008	0.005	0.000	0.010	0.003	0.007	-0.011

*Upfield shifts Δδ in ppm, are taken with positive sign.

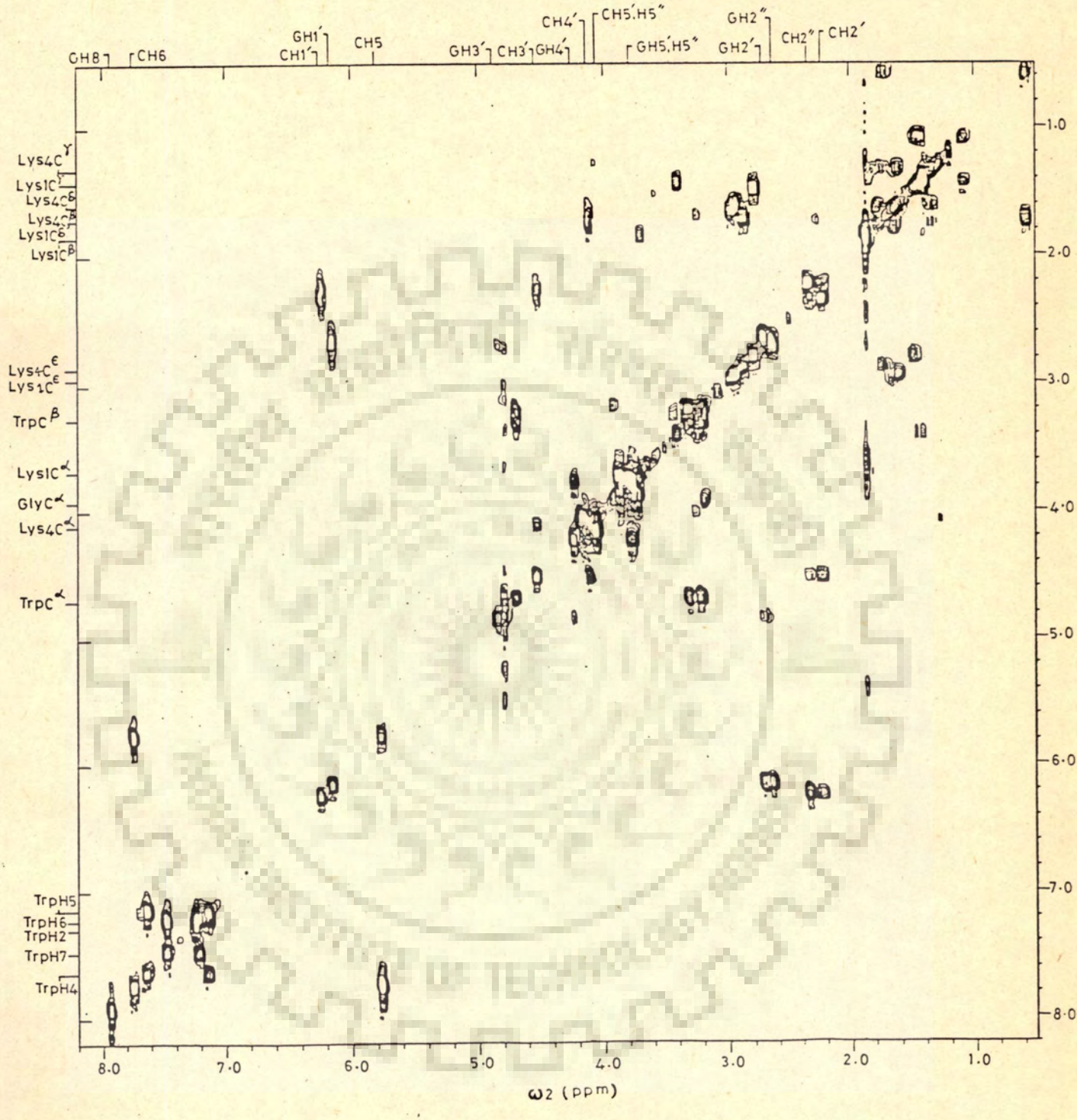


Fig. 5.8 500 MHz COSY spectrum of a mixture of 11 mM d-GpC and 5.5 mM Lys-Trp-Gly-Lys OtBu at 297 K for the sample used in Fig. 5.6.

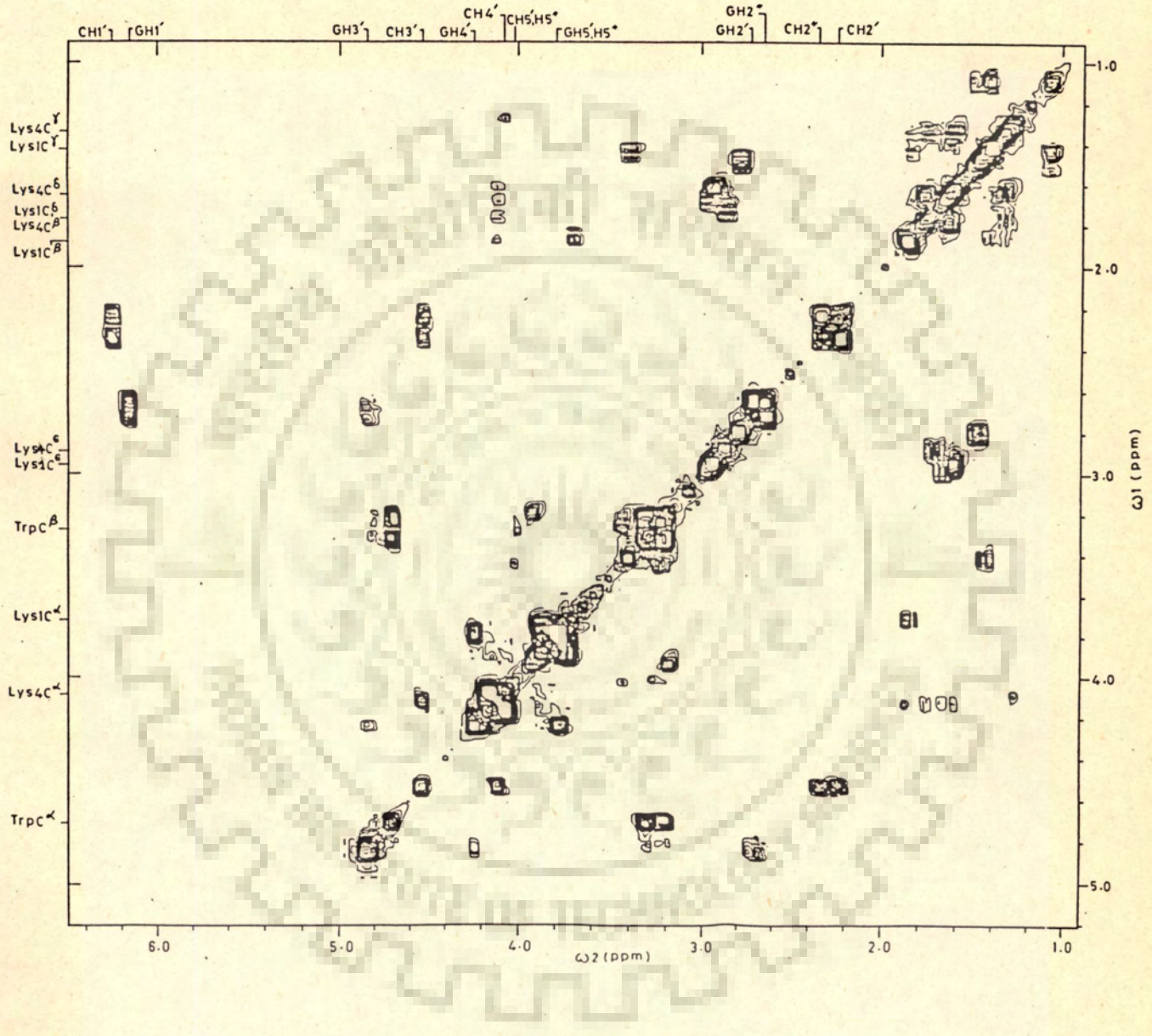


Fig. 5.8 (a) Portion of COSY spectrum of Fig. 5.8 in an expanded scale

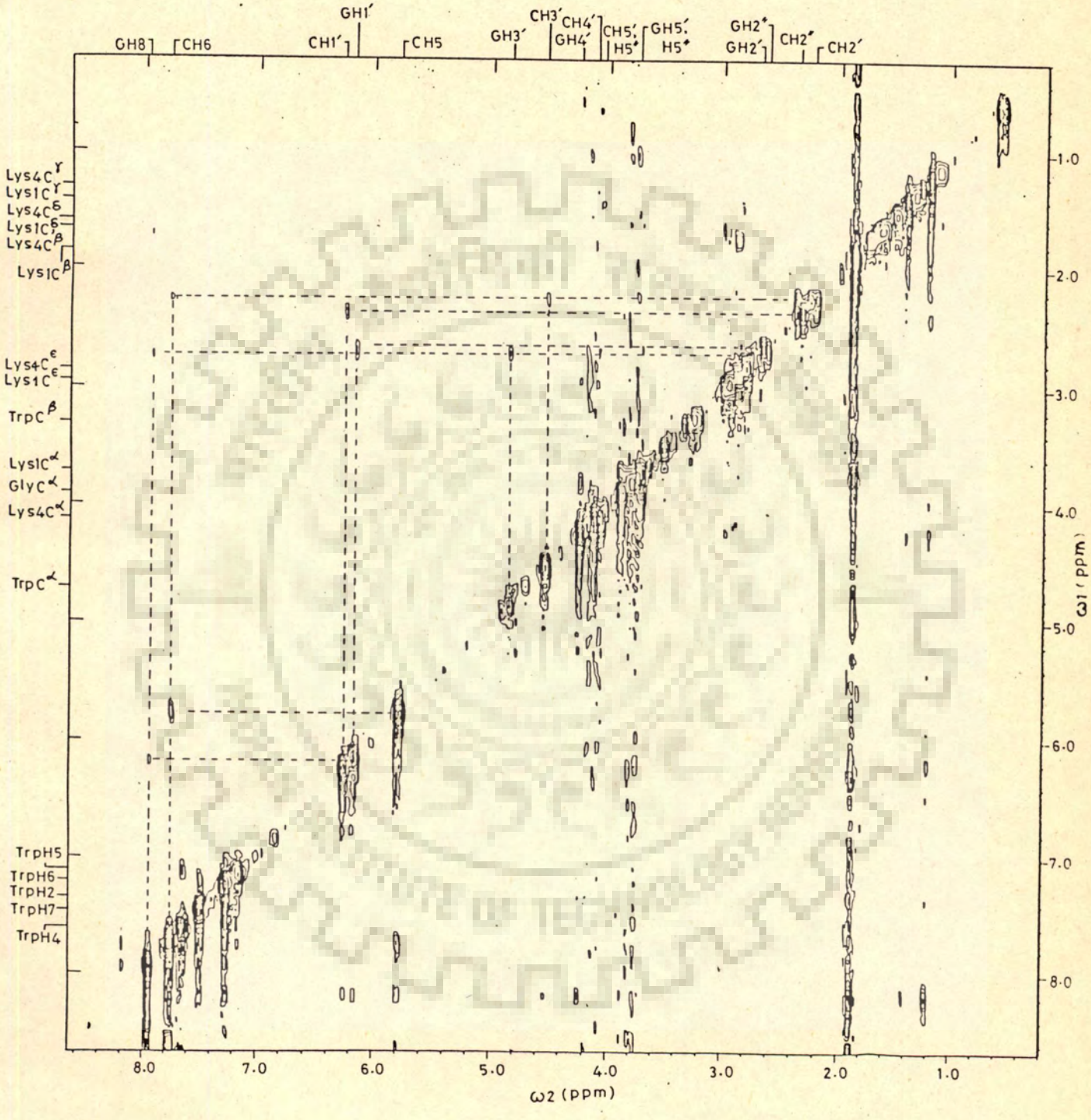


Fig. 5.9 500 MHz NOESY spectrum of a mixture of 11 mM d-GpC and 5.5mM Lys-Trp-Gly-Lys OtBu at 297 K for the same sample used in Fig.5.6.

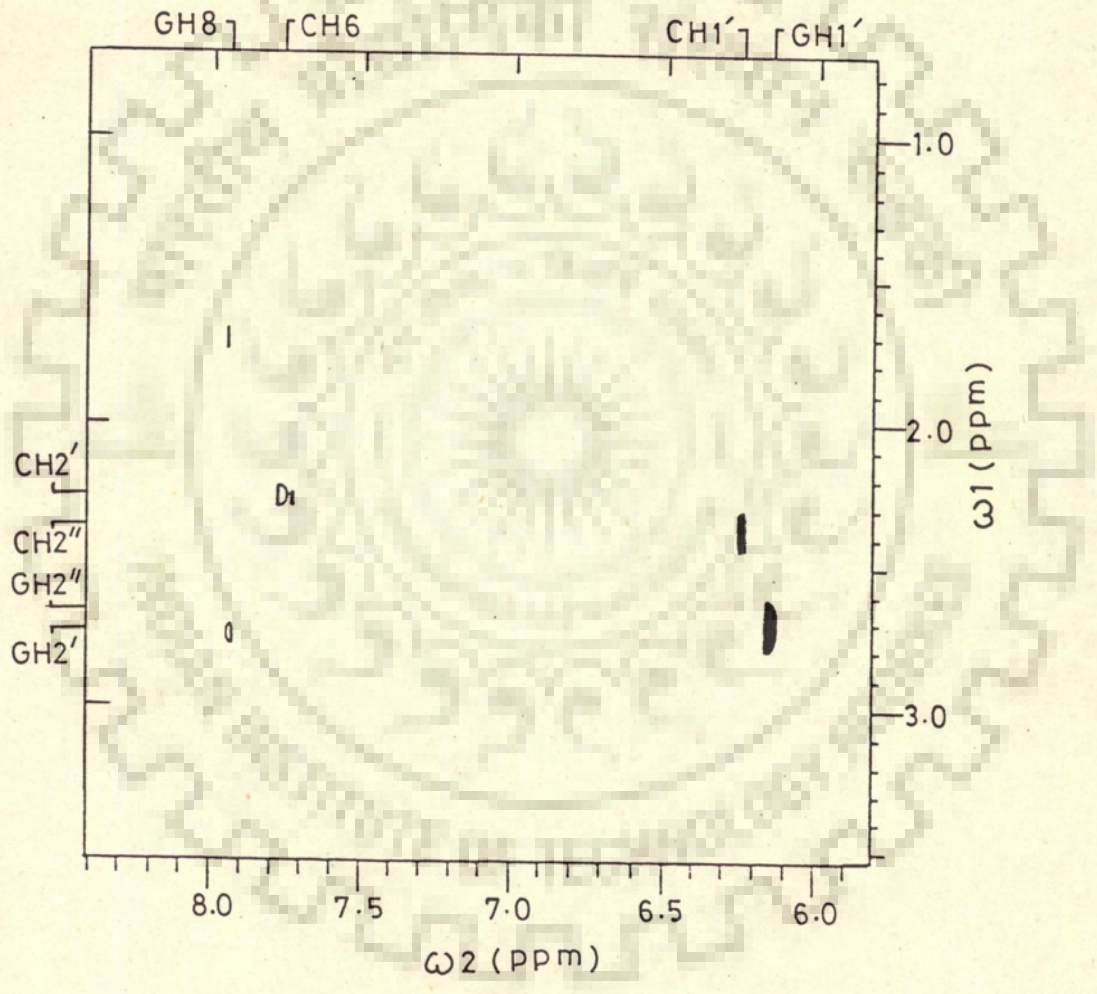


Fig. 5.9 (a) Portion of NOESY spectrum of Fig. 5.9 in an expanded scale

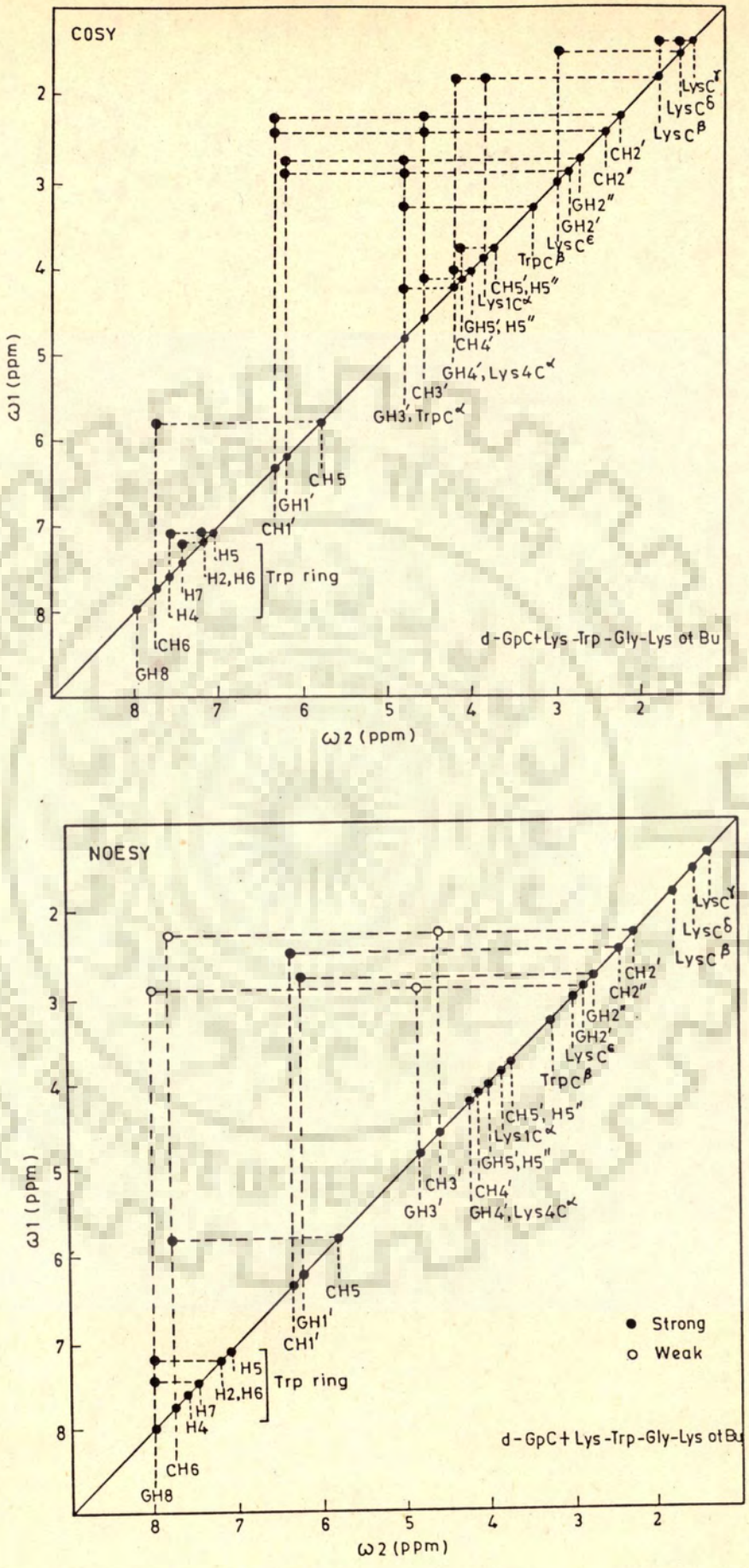


Fig. 5.10 Schematic representation of the results of COSY and NOESY spectra of a mixture of d-GpC and Lys-Trp-Gly-Lys OtBu shown in figs. 5.8 and 5.9

The relative intensities of NOE cross peaks in d-GpC bound to Lys-Trp-Gly-Lys OtBu are as follows:

GH8 ... GH2'	intra-residue	weak	(1)
GH8 ... GH2"	"	not seen	(2)
GH8 ... GH1'	"	"	(3)
CH6 ... GH1'	inter-residue	"	(4)
CH6 ... GH2"	"	"	(5)
CH6 ... CH2'	intra-residue	weak	(6)
CH6 ... CH2"	"	not seen	(7)
CH6 ... CH1'	"	not seen	(8)
GH1'... GH2'	"	"	(9)
GH1'... GH2"	"	strong	(10)
CH1'... CH2'	"	not seen	(11)
CH1'... CH2"	"	strong	(12)
GH8 ... Trp H7,H2	inter-molecular	strong	(13)

For both G and C the base to H2' connectivity, (1) and (6), are stronger than that of base to H2", (2) and (7), so that the glycosidic bond rotation is inferred to be in anti conformation for both bases. There is no NOE connectivity which could demonstrate that d-GpC mixture exists as double—helix. The presence of intra-residue base-H2' connectivities i.e. (1) and (6) also demonstrate presence of right—handed DNA. The inter-residue connectivities

i.e. (4) and (5) are not seen. This may be so due to the fact that d-GpC mini double helix is too small and is not expected to follow geometries as in standard B—DNA. The presence of inter-molecular peak (13), i.e. between base proton and Tryptophan ring protons provides direct evidence that Trp stacks with bases in the complex of d-GpC with Lys-Trp-Gly-Lys OtBu.

BINDING OF d-GpC TO Lys-Tyr-Lys

Figures 5.11 and 5.12 show the results of binding of d-GpC to tripeptide Lys-Tyr-Lys at different temperatures. Changes in chemical shift of all protons is given in Table 5.3. It is found that $T_{\frac{1}{2}}$ value on binding increases from 286K to 302 - 307 K. Tyr ring protons shift upfield by about 0.028 ppm at 285 K which is more than the corresponding change observed (0.014 ppm) on the binding to d-CpG. This change decreases with temperature and, therefore, the binding seems to be preferential to double-helix. The upfield shift is due to stacking/intercalation between base pairs of d-GpC.

Figures 5.13 and 5.14 show results of COSY and NOESY spectra of a mixture of d-GpC with Lys-Tyr-Lys for the same sample solution as above. Figures 5.13a, b and 5.14a, b are portions of the same spectra on expanded scale and Figure 5.15 is the schematic representation of 2D-NMR results. Once again it is found from COSY spectra that all the sugar

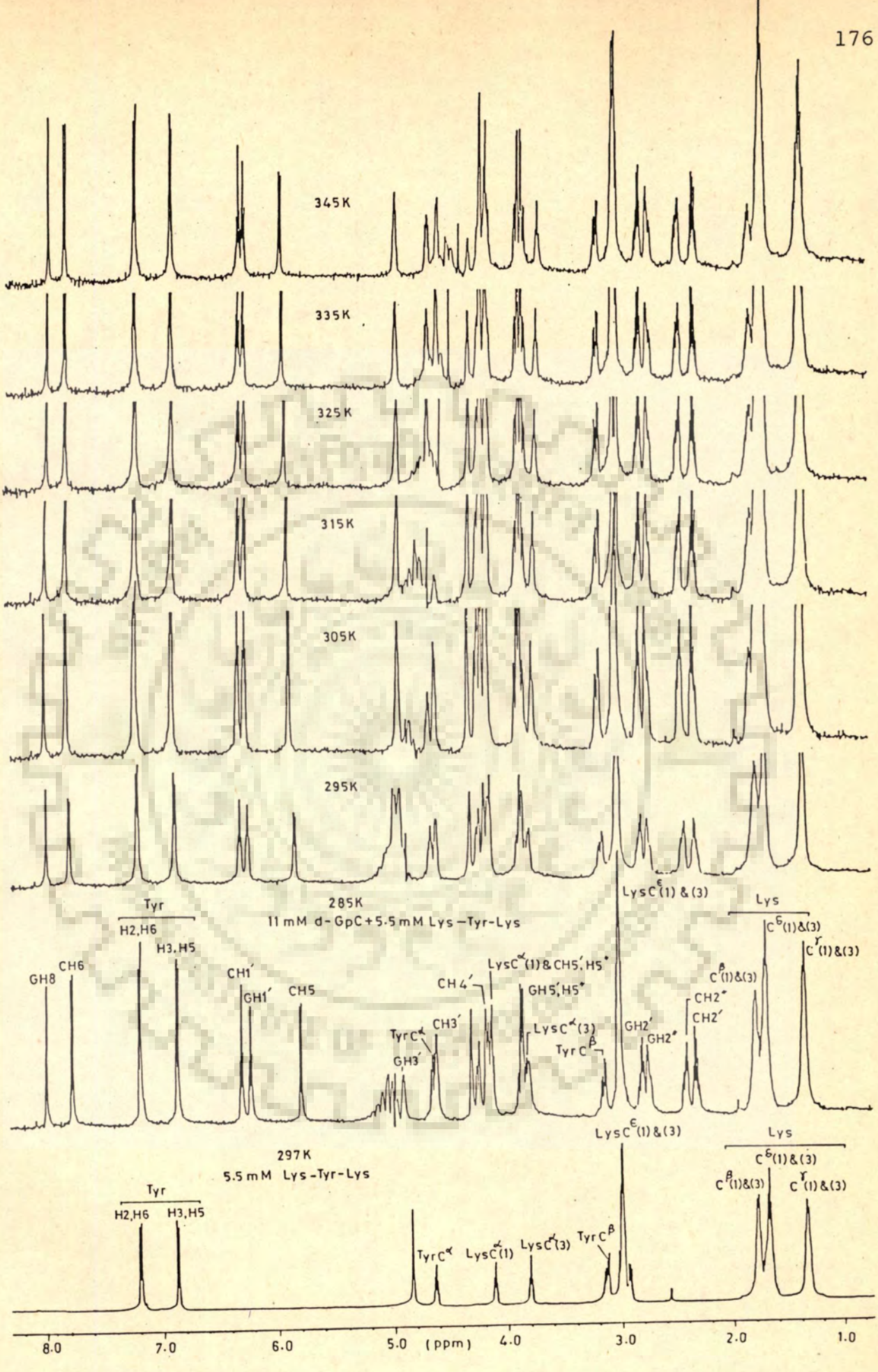


Fig. 5.11 500 MHz proton NMR spectra of a mixture of 11 mM d-GpC and 5.5 mM Lys-Tyr-Lys at different temperatures in D_2O (pH=7.0) containing 0.25 mM EDTA. 500 MHz spectrum of 5.5 mM Lys-Tyr-Lys in D_2O at 297 K is also shown below. Ref. DSS.

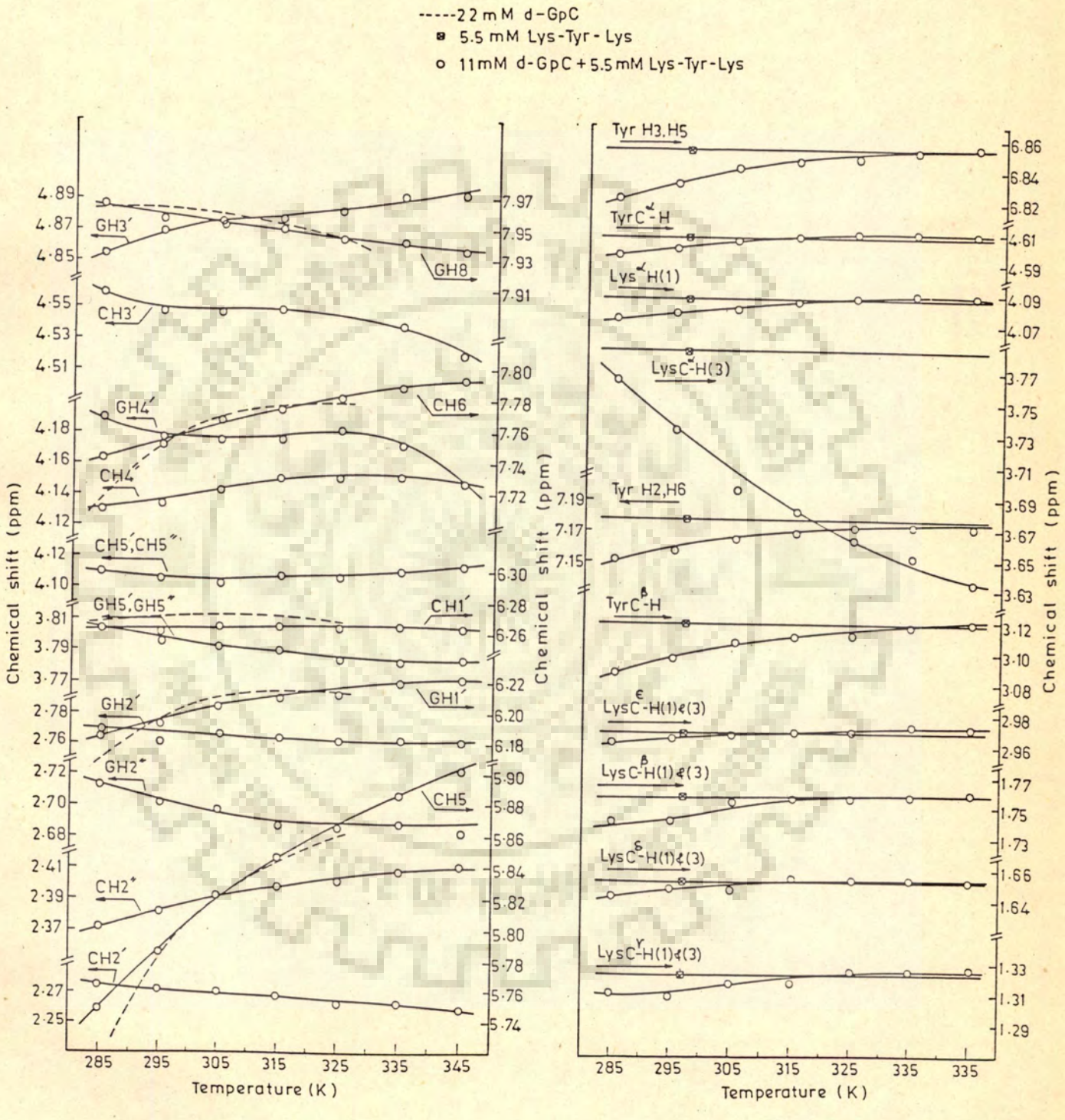


Fig. 5.12 Chemical shifts of some of the nucleotide and peptide protons as a function of temperature in uncomplexed nucleotide/peptide and in complexed state taken from data in Fig. 5.11.

Table 5.3 : Changes in chemical shift, $\Delta\delta$ in ppm, of peptide and nucleotide protons on binding of 11 mM d-GpC to 5.5 mM Lys-Tyr-Lys at indicated temperatures

SYSTEM	Temp. (K)	GH8	CH6	CH5	CH1'	GH1'	CH2'	GH2"	GH2'	GH2"
d-GpC	285	7.962	7.717	5.699	6.264	6.161	2.294	2.361	2.732	2.732
	297	7.960	7.756	5.793	6.268	6.197	2.284	2.373	2.754	2.703
	315	7.950	7.772	5.842	6.266	6.210	2.284	2.383	2.758	2.694
d-GpC + Lys-Tyr-Lys	285	7.965	7.743	5.749	6.263	6.184	2.275	2.373	2.769	2.712
	207	7.956	7.753	5.767	6.260	6.193	2.272	2.383	2.761	2.702
	315	7.950	7.775	5.846	6.265	6.210	2.267	2.397	2.764	2.687
$\Delta\delta$ (ppm)*	285	-0.003	-0.026	-0.050	0.001	-0.023	0.019	-0.012	-0.037	0.020
	297	0.004	0.003	0.026	0.008	0.004	0.012	-0.010	-0.007	0.001
	315	0.000	-0.003	-0.004	0.001	0.000	0.017	-0.014	-0.006	0.007

SYSTEM	Temp. (K)	CH3'	GH3'	CH4'	GH4'	CH5'/H5"	GH5'/H5"	Tyr H2,H6	Tyr H3,H5
d-GpC / Lys-Tyr-Lys	285	4.560	-	4.124	4.193	4.096	3.808	-	-
	297	4.558	-	4.141	4.181	4.097	3.810	7.177	6.854
	315	-	4.876	4.144	4.176	4.093	3.809	-	-
d-GpC + Lys-Tyr-Lys	285	4.558	-	4.130	4.189	4.109	3.803	7.151	6.823
	297	4.547	-	4.134	4.177	4.106	3.795	7.156	6.830
	315	-	4.868	4.151	4.174	4.108	3.790	7.168	6.847
$\Delta\delta$ (ppm)*	285	0.002	-	-0.006	0.004	-0.013	0.005	0.026	0.031
	297	0.011	-	0.011	0.004	-0.009	0.015	0.021	0.024
	315	-	0.008	-0.007	0.002	-0.015	0.019	0.009	0.007

Contd.

SYSTEM	Temp. (K)	Tyr C^{α}-H	Tyr C^{β}-H	Lys C^{α}-H (1)	Lys C^{α}-H (3)	Lys C^{β}-H (1)&(3)	Lys C^{γ}-H (1)&(3)	Lys C^{δ}-H (1)&(3)	Lys C^{ε}-H (1)&(3)
Lys-Tyr-Lys	297	4.608	3.121	4.089	3.785	1.758	1.323	1.655	2.971
d-GpC + Lys-Tyr-Lys	285	4.597	3.089	4.075	3.766	1.742	1.313	1.644	2.965
	297	4.601	3.099	4.079	3.734	1.743	1.309	1.649	2.967
	315	4.608	3.111	4.087	3.682	1.756	1.317	1.656	2.970
$\Delta\delta$ (ppm)*	285	0.011	0.032	0.014	0.019	0.016	0.010	0.011	0.006
	297	0.007	0.022	0.010	0.051	0.015	0.014	0.006	0.004
	315	0.000	0.010	0.002	0.103	0.002	0.006	-0.001	0.001

*Upfield shifts, $\Delta\delta$ in ppm, are taken with positive sign.

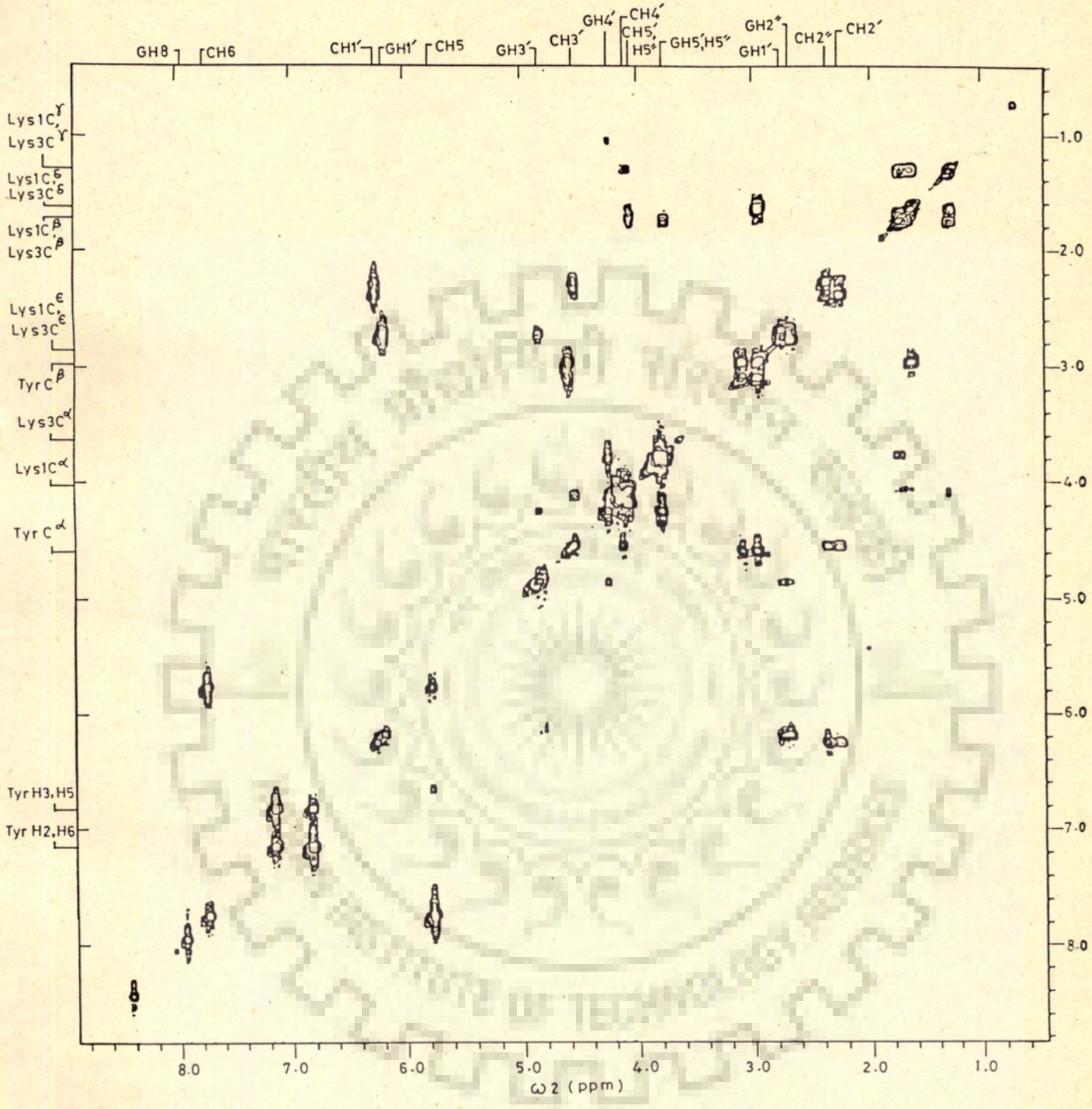


Fig. 5.13 500 MHz COSY spectrum of a mixture of 11 mM d-GpC and 5.5mM Lys-Tyr-Lys in D_2O (pH=7.0) containing 0.25 mM EDTA at 297 K

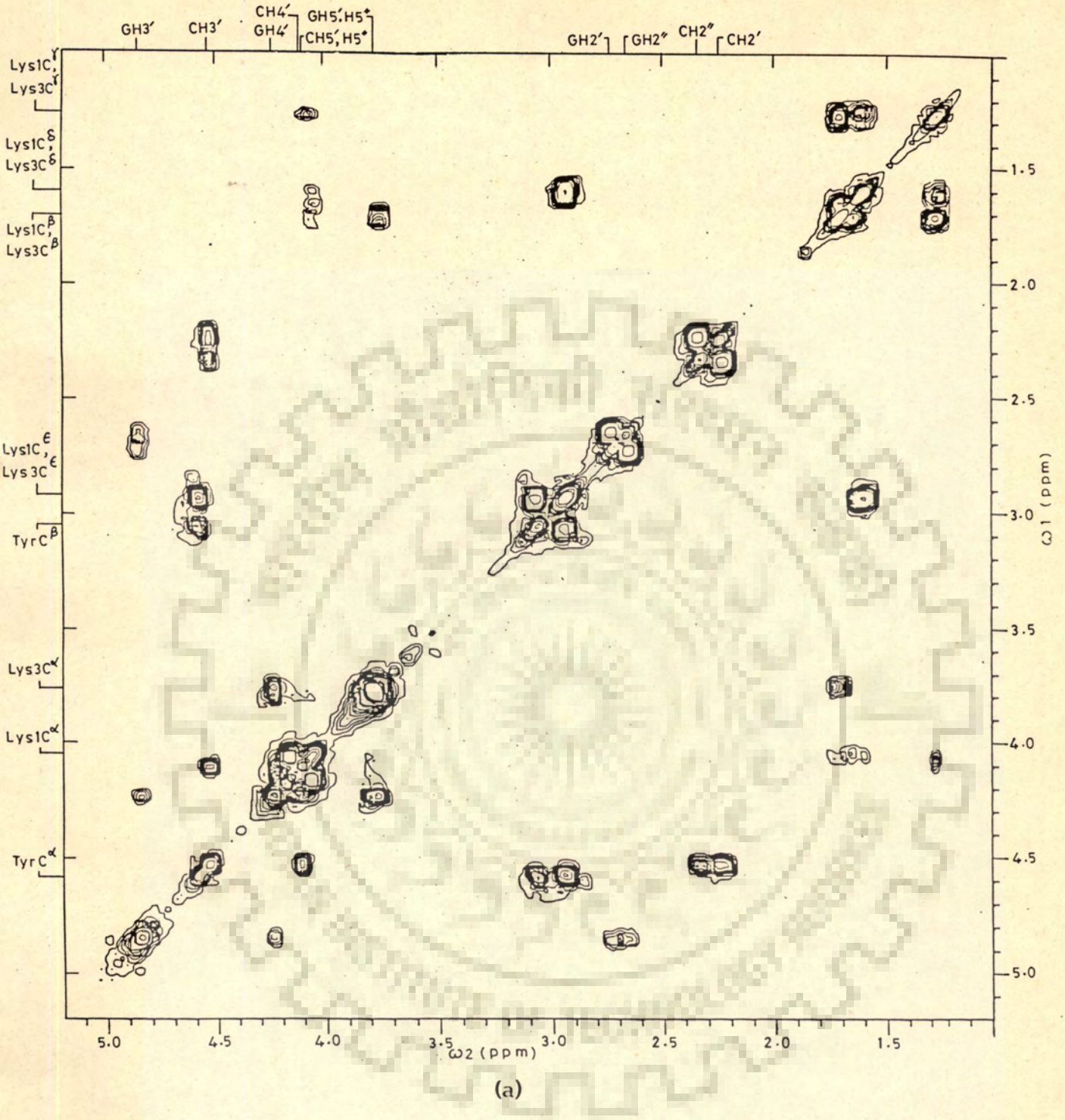
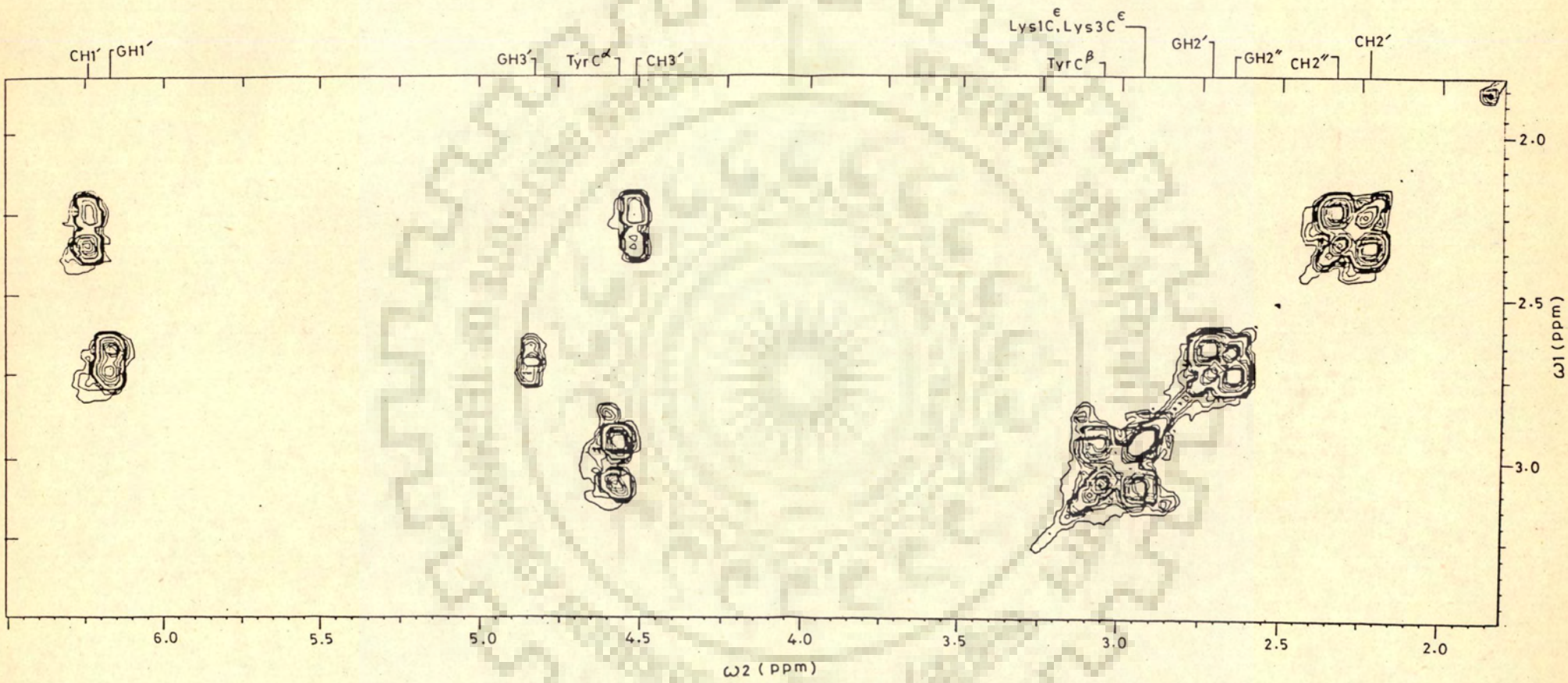


Fig. 5.13 (a), (b) Portions of COSY spectra of fig. 5.13 in an expanded scale



(b)

Fig. 5.13 Contd.

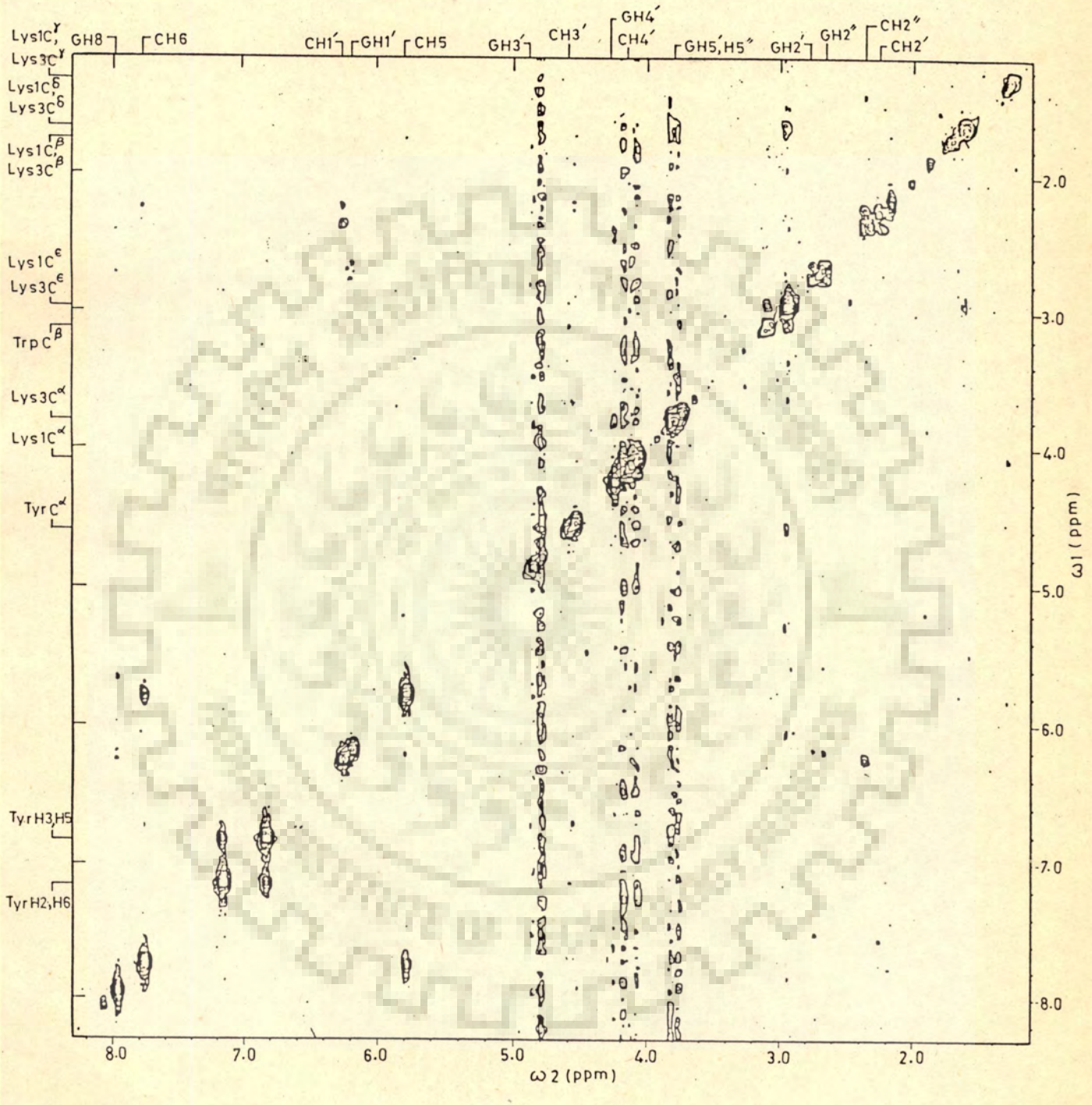


Fig. 5.14 500 MHz NOESY spectrum of a mixture of 11 mM d-GpC and 5.5mM Lys-Tyr-Lys in D_2O (pH=7.0) containing 0.25 mM EDTA at 297 K

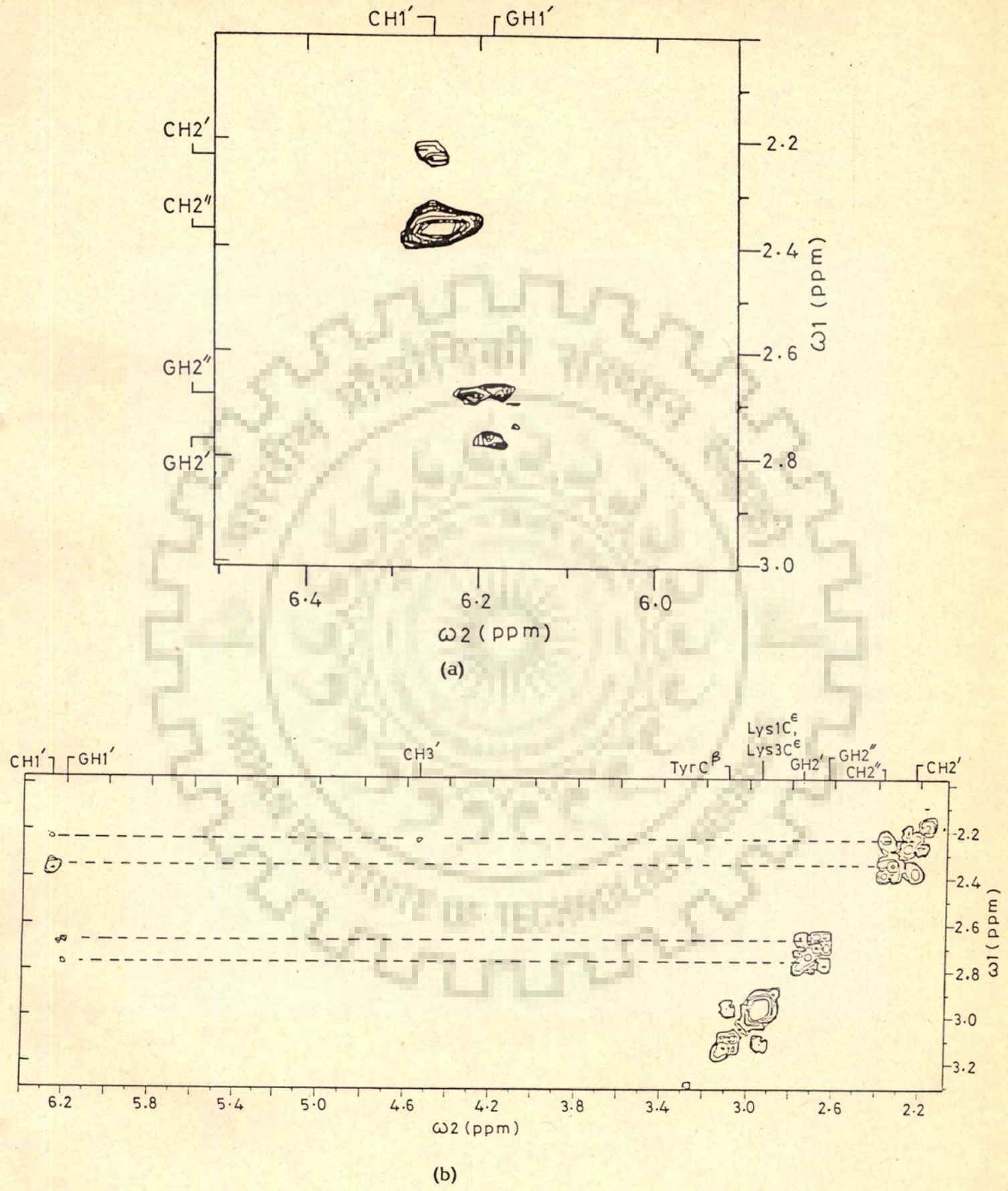


Fig. 5.14 (a), (b) Portions of NOESY spectra of fig. 5.14 in an expanded scale

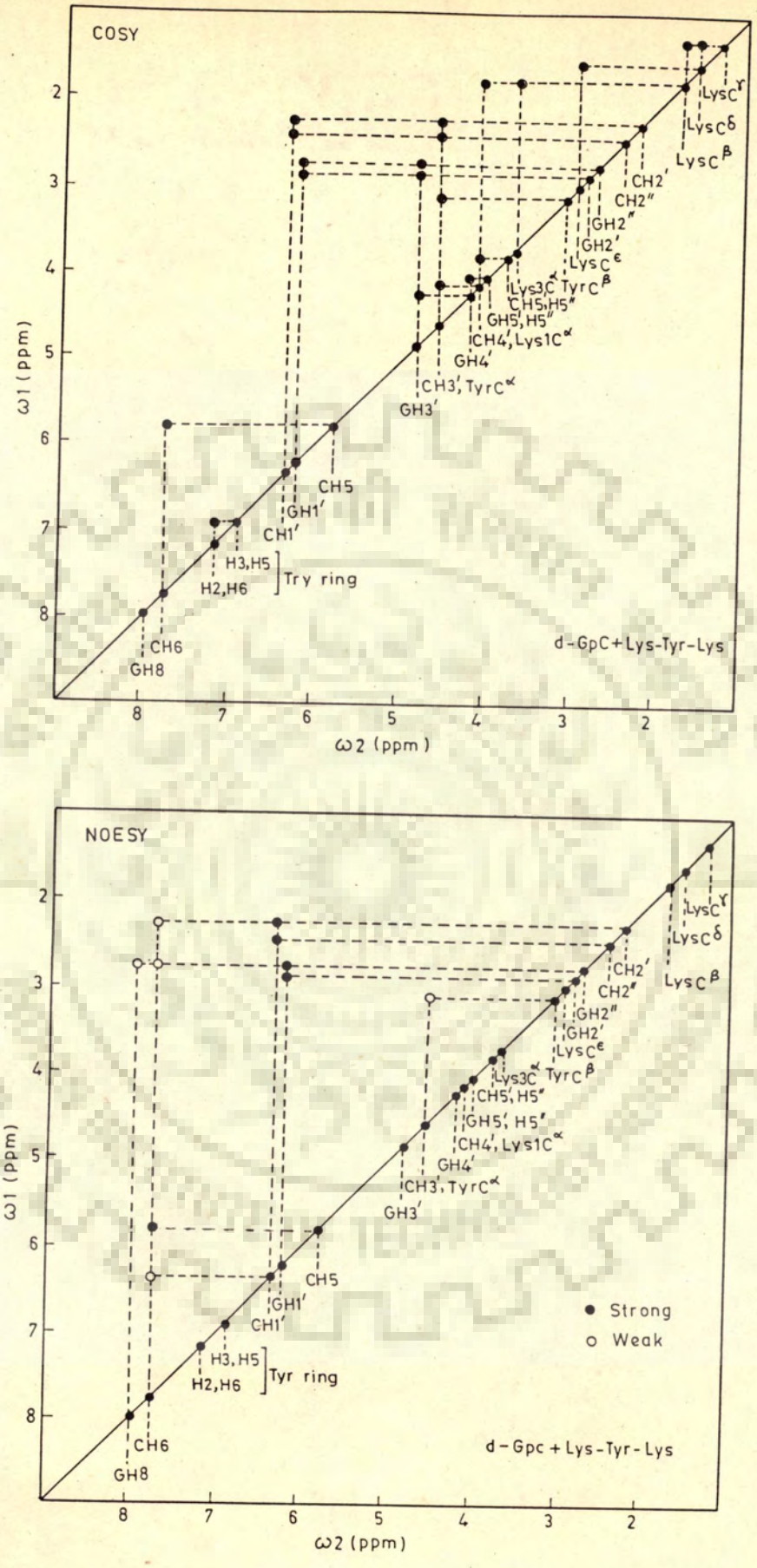


Fig. 5.15 Schematic representation of the results of COSY and NOESY spectra of a mixture of d-GpC and Lys-Tyr-Lys shown in figs. 5.13 and 5.14

proton J-connectivities are the same as that in uncomplexed d-GpC and, therefore both the bases have sugar conformation as 01' endo.

The relative intensities of NOE cross peaks in the complex of d-GpC with Lys-Tyr-Lys are as follows:

GH8 ... GH2'	intra-residue	Not seen	(1)
GH8 ... GH2"	"	weak	(2)
GH8 ... GH1'	"	not seen	(3)
CH6 ... GH1'	inter-residue	"	(4)
CH6 ... GH2"	"	"	(5)
CH6 ... CH2'	intra-residue	weak	(6)
CH6 ... CH2"	"	not seen	(7)
CH6 ... CH1'	"	not seen	(8)
GH1' ... GH2'	"	strong	(9)
GH1'... GH2"	"	strong	(10)
CH1'... CH2'	"	strong	(11)
CH1'... CH2"	"	strong	(12)
CH3'... Tyr C^{β}	Inter-molecular	strong	(13)

Once again inter-residue peaks (4) and (5) are not seen. Also there is no NOE between protons on one strand and the protons on other strand. As in complex of d-GpC with Lys-Trp-Gly-Lys OtBu, the complex with Lys-Tyr-Lys also

forms a mini double-helix and is not expected to follow stranded geometries of BDNA. Appearance of intra-residue peak nos. (1) and (6), i.e. stronger than corresponding base H2" NOE'S indicates that χ_{CN} is in anti conformation for both G and C. Further these intra-residue NOE'S ascertain the presence of right-handed B-DNA. The existence of inter-molecular peak No. 13, that is, between sugar proton CH3' and Tyrosine C^{β} proton indicates proximity of Tyr residue to the dinucleotide.

The conformation of the complex d-GpC with Lys-Trp-Gly-Lys and Lys-Tyr-Lys has all sugars with O1'-endo pucker, all χ_{CN} to be anti and helix sense as right-handed. The Trp ring stacks with bases in the complex of d-GpC with Lys-Trp-Gly-Lys while in complex of GpC with Lys-Tyr-Lys, the proximity of dinucleotide to tripeptide is observed.

CHAPTER - 6

d-CpCpGpG AND ITS BINDING TO TRIPEPTIDE Lys-Tyr-Lys

500 MHz proton NMR spectra of deoxytetranucleotide d-CpCpGpG and its mixture with tripeptide Lys-Tyr-Lys as a function of temperture in the range 285 - 350 K were studied. 2D-COSY and NOESY spectra were taken to find the conformation of d-CpCpGpG alone and its complex with tripeptide.

d-CpCpGpG

A. Assignments

The bases in d-CpCpGpG have been numbered as follows:

d - C p C p G p G
1 2 3 4
ext. int. int. ext.

Figure 6.1 shows the spectra of d-CpCpGpG in D_2O solution as a function of temperature and Table 6.1 shows the chemical shift values for the same. It is easy to distinguish non-exchangeable base protons [109] from their positions in spectra. GH8 as expected resonates at ~7.89 ppm at 297 K, the two CH6 protons appear at 8.01 and 7.57 ppm while the

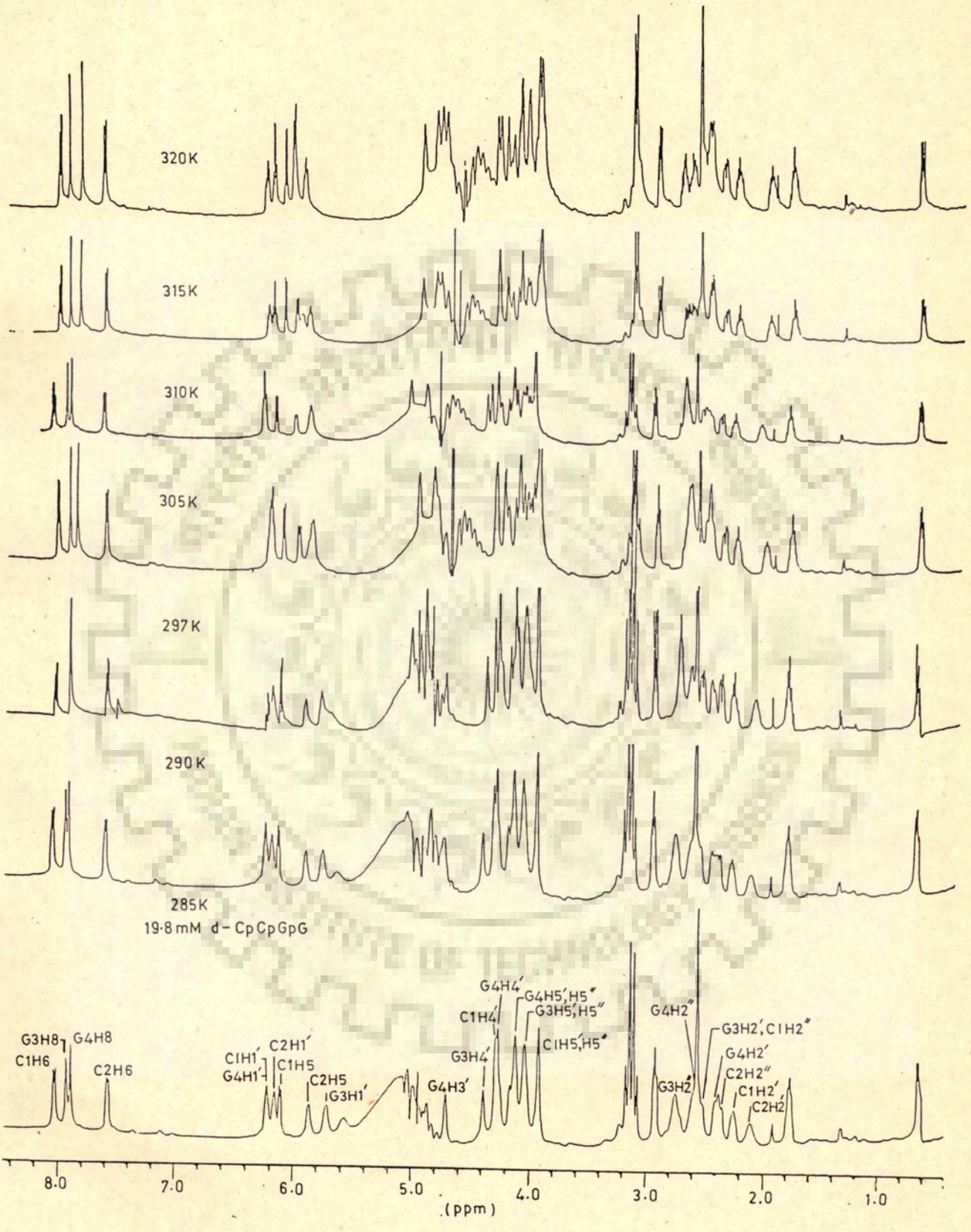


Fig. 6.1 500 MHz proton NMR spectra of deoxytetranucleotide 19.8 mM d-CpCpGpG in D_2O (pH=7.0) solution containing 0.25 mM EDTA at indicated temperatures. Ref. DSS.

Table 6.1 : Chemical shift values of various protons at 500 MHz for 19.8 mM d-CpCpGpG in D_2O (pH=7.0) in temperature range 285K-320K.

Temp. (K)	C1H6	C2H6	G3H8	G4H8	C1H5	C2H5	C1H1'	C2H1'	G3H1'	G4H1'
285	8.026	7.575	7.926	7.889	6.102	5.863	6.218	6.156	5.720	5.579
290	8.016	7.562	7.908	7.881	6.099	-	6.210	6.157	-	-
297	8.013	7.566	7.887	7.887	6.097	5.896	6.206	6.168	5.755	5.704
305	8.010	7.580	7.862	7.889	6.100	5.933	6.197	6.197	5.808	5.808
310	8.010	7.593	7.846	7.908	6.092	5.961	6.194	6.207	5.848	5.848
315	8.011	7.610	7.833	7.920	6.092	5.994	6.190	6.233	5.882	5.953
320	8.012	7.627	7.824	7.932	6.090	6.020	6.188	6.249	5.925	6.020

Temp. (K)	C1H2'	C1H2"/ G3H2'	G3H2"	C2H2'	C2H2"	G4H2'	G4H2"	C1H3'	C2H3'	G3H3'
285	2.292	2.522	2.745	2.100	2.359	2.406	2.554	4.872	-	5.025
290	-	2.504	2.726	-	-	-	-	4.865	4.791	5.010
297	2.228	2.494	2.692	2.047	2.342	2.412	2.597	4.843	4.785	4.986
305	2.224	2.477	2.667	2.002	2.337	2.433	2.643	4.824	4.774	4.957
310	2.223	2.470	2.634	1.977	2.336	2.456	2.646	4.815	4.774	4.940
315	2.226	2.456	2.613	1.958	2.338	2.477	2.668	4.806	4.769	4.924
320	2.226	2.446	2.604	1.945	2.343	2.478	2.690	4.799	4.756	4.911

Temp. (K)	G4H3'	C1H4'	G3H4'	G4H4'	C1H5'/H5"	G3H5'/H5"	G4H5/H5"
285	4.702	4.279	4.375	4.263	3.918	4.040	4.116
290	4.693	4.277	4.365	4.253	3.916	4.033	4.109
297	4.713	4.277	4.344	4.234	3.918	4.024	4.103
305	4.712	4.279	4.315	4.213	3.924	4.022	4.015
310	4.715	4.284	4.284	4.198	3.927	4.023	4.092
315	4.719	4.281	4.281	4.185	3.928	4.024	4.091
320	4.720	4.288	4.262	4.177	3.928	4.025	4.091

two CH5's resonate at 6.1 and 5.6 ppm. From the observed change in chemical shift on increasing temperature from 285 to 320 K, it is possible to distinguish between C2H6 (internal) and C1H6 (external) protons. However, complete unambiguous assignment of each and every resonance is possible only by 2D-COSY and NOESY spectra. The chemical shifts versus temperature plots (Figure 6.2) monitor the helix coil transition from a stable duplex at low temperatures to unstacked strands at high temperatures. The $T_{\frac{1}{2}}$ value of helix coil transition for some of the base and sugar protons are as follows:

	$T_{\frac{1}{2}}$ (our results)	$T_{\frac{1}{2}}$ (Patel)
G4H8	310 K	310 K
C2H6	310 K	326 K
C2H1'	312 K	-
G3H1'	310 K	-
C2H5	310 K	315 K

The observed values are lower than that observed by Patel [109].

Figures 6.3 and 6.4 show 500 MHz two dimensional COSY and NOESY spectra of d-CpCpGpG at 297 K while Figs.6.3a, b, c, d and 6.4a, b, c are protons of these spectra expanded to show specific connectivities. The sample solution

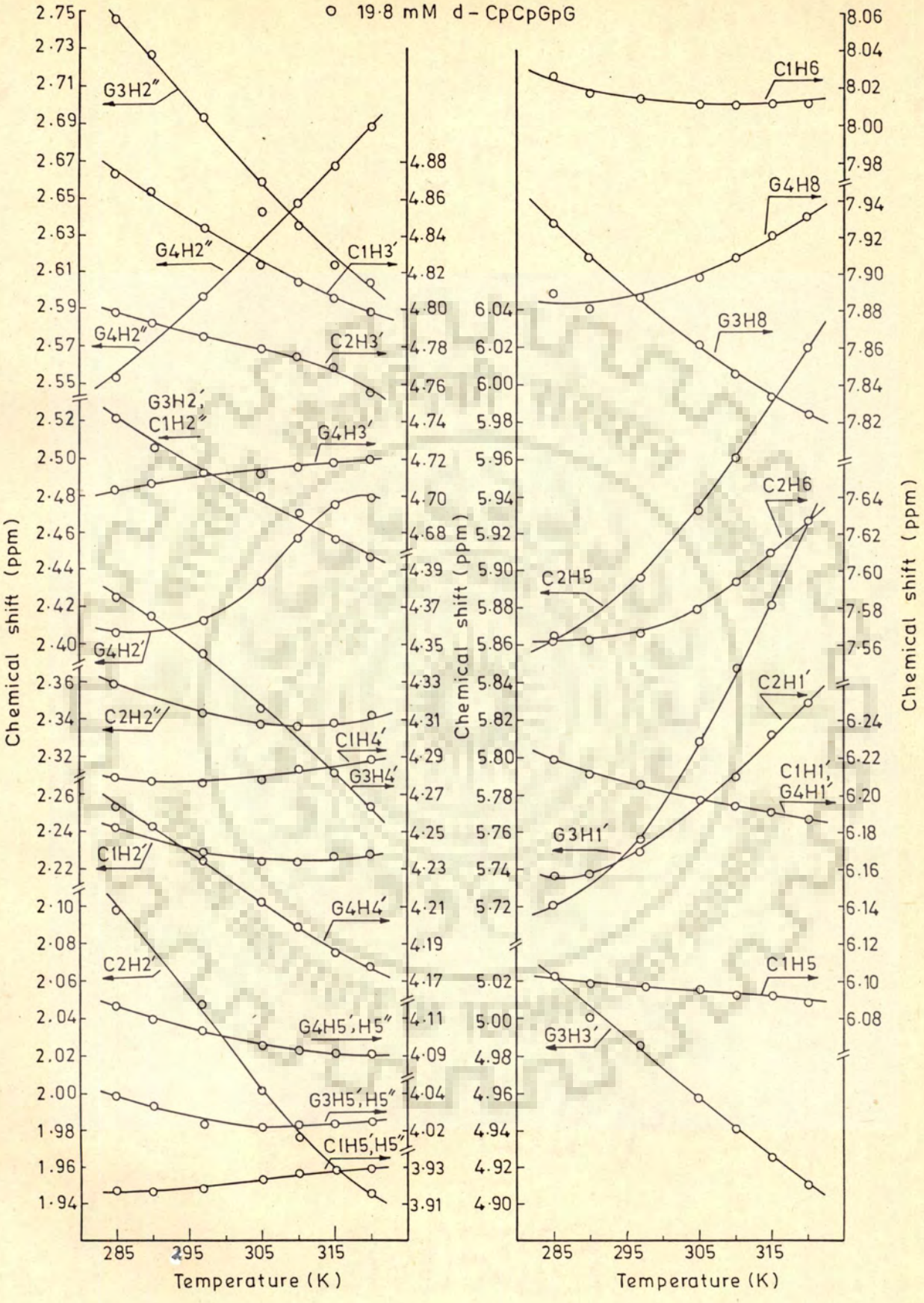


Fig. 6.2 Chemical shfit of various protons of d-CpCpGpG as a function of temperature taken from spectra of fig. 6.2

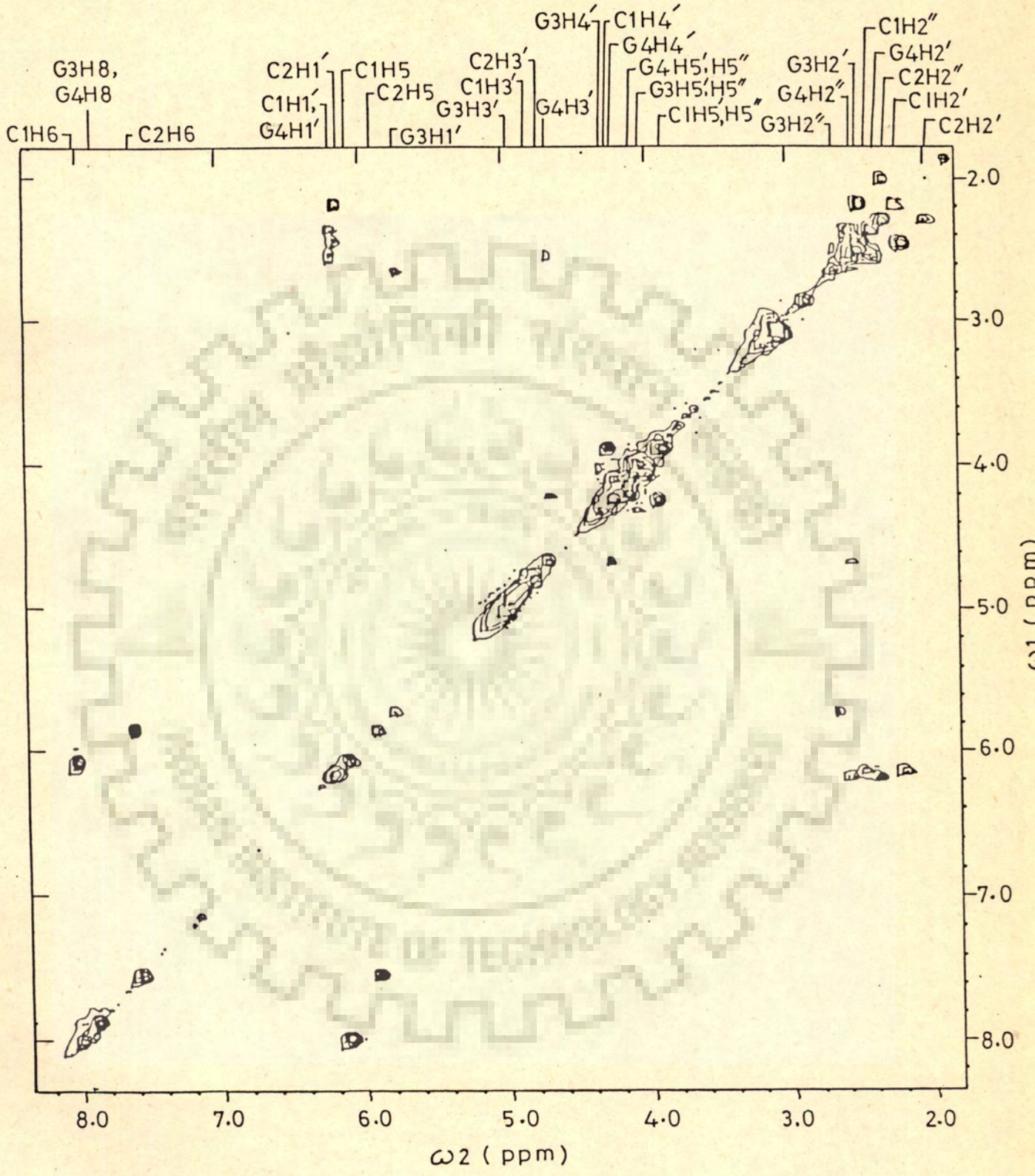


Fig. 6.3 500 MHz COSY spectrum of 19.8 mM d-CpCpGpG solution in D_2O containing 0.25 mM EDTA (pH=7.0) at 297 K

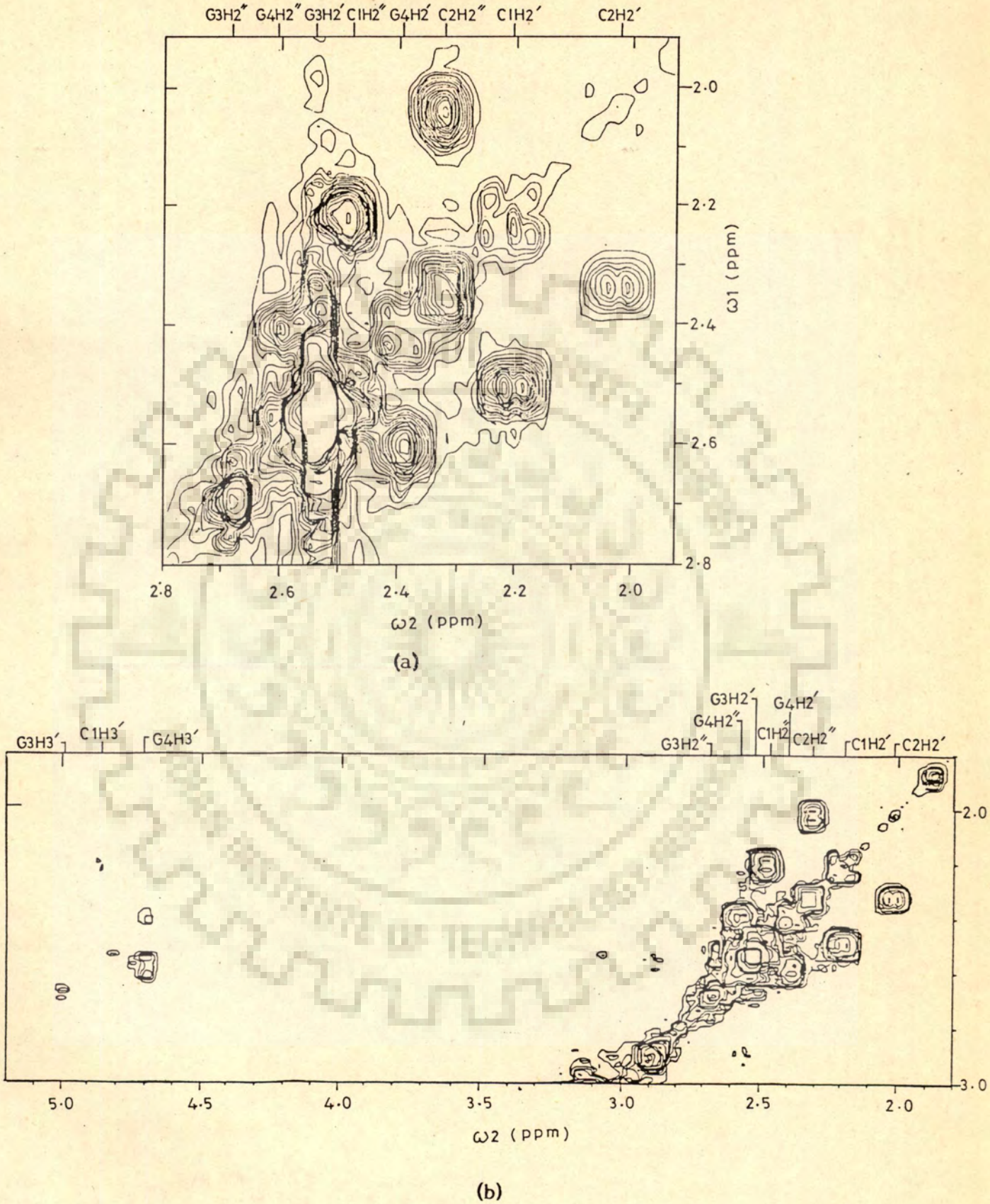
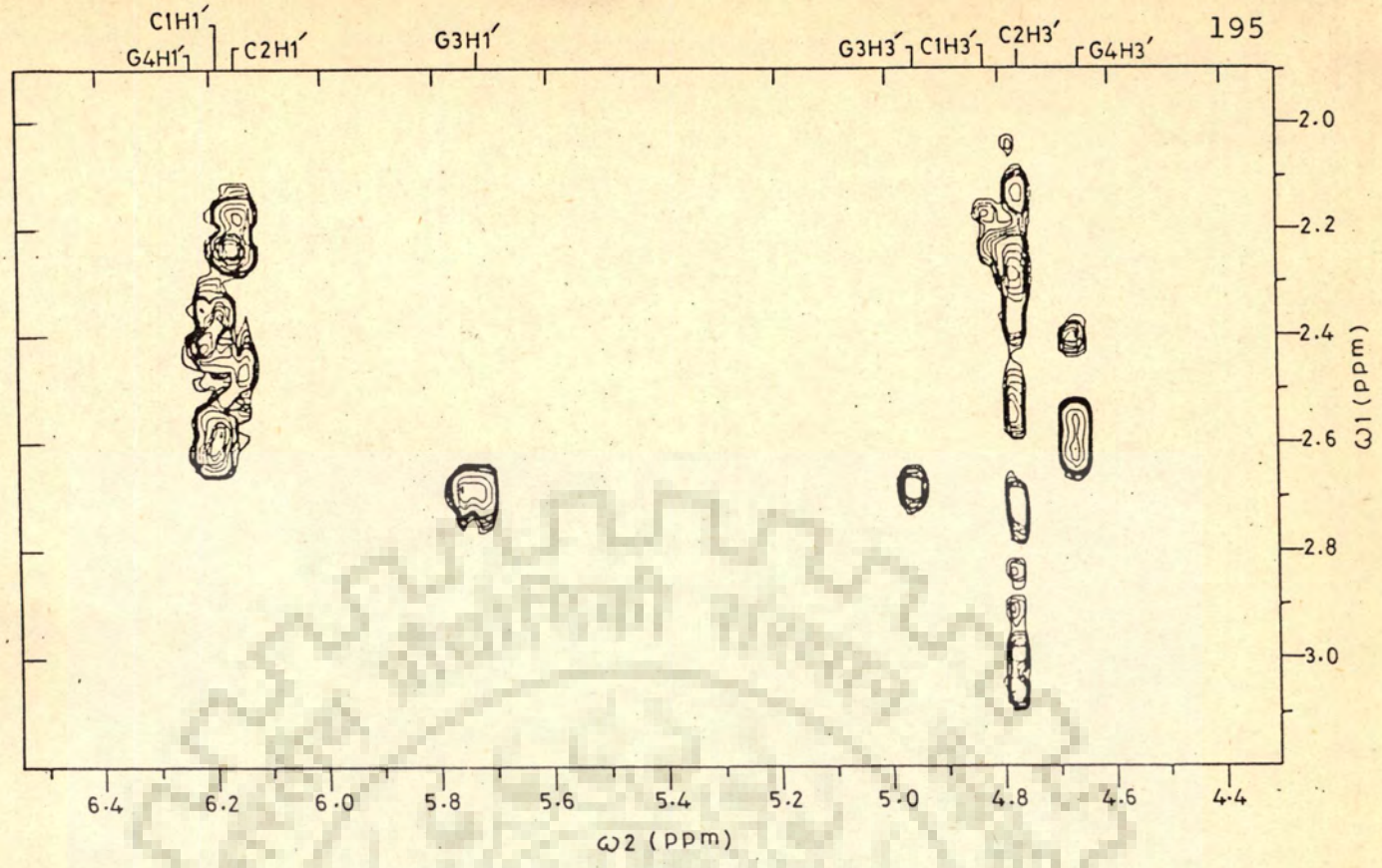
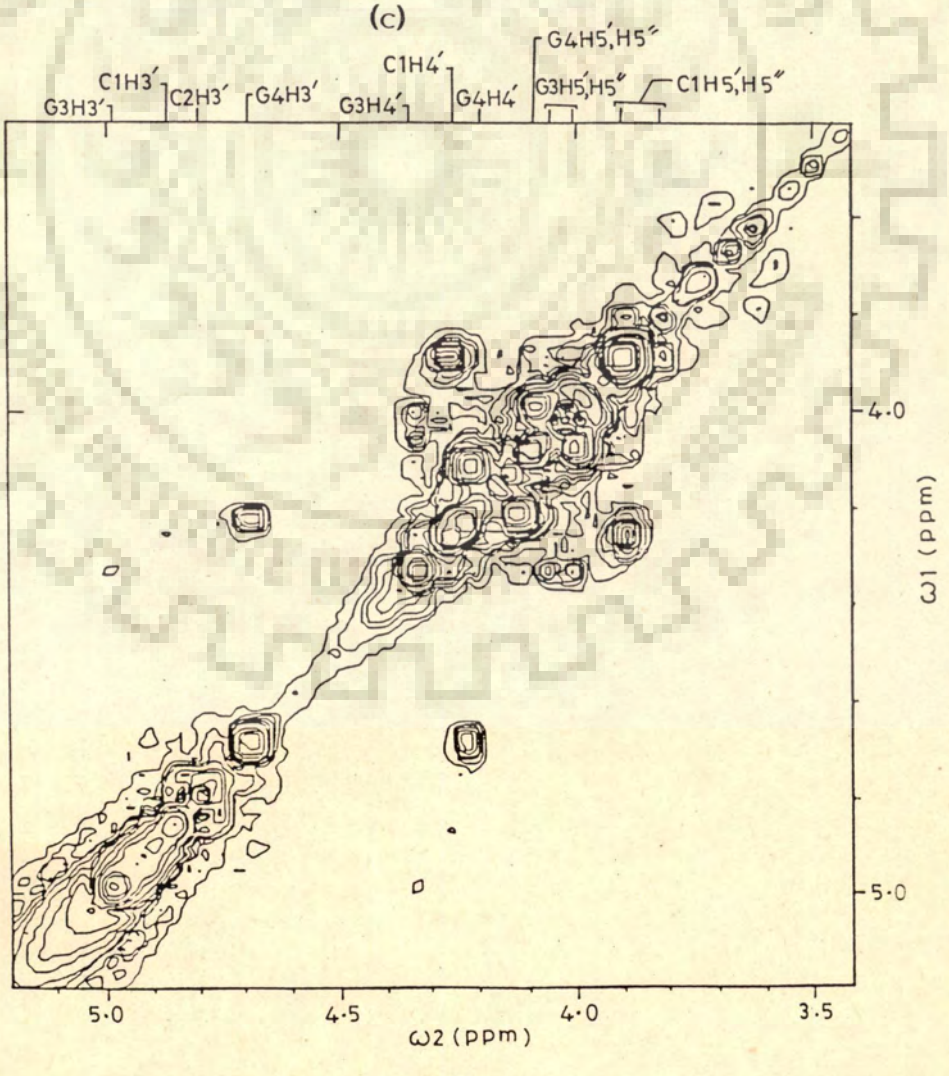


Fig. 6.3 **(a), (b), (c)** and **(d)** Portions of COSY spectra of fig. 6.3 shown in expanded scales.



(c)



(d)

Fig. 6.3 Contd.

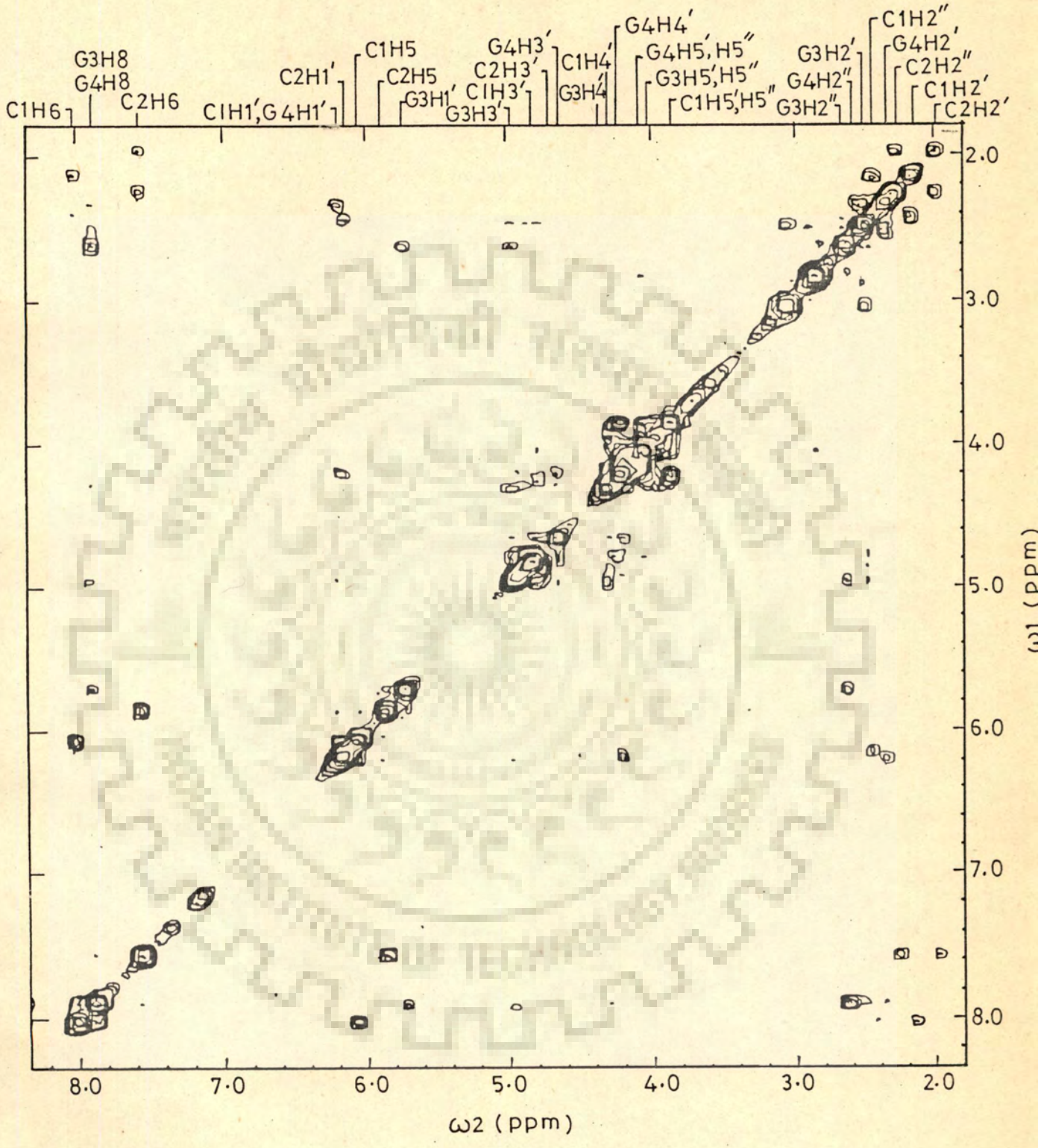


Fig. 6.4 500 MHz NOESY spectrum of 19.8 mM d-CpCpGpG solution in D_2O containing 0.25 mM EDTA (pH=7.0) at 297 K

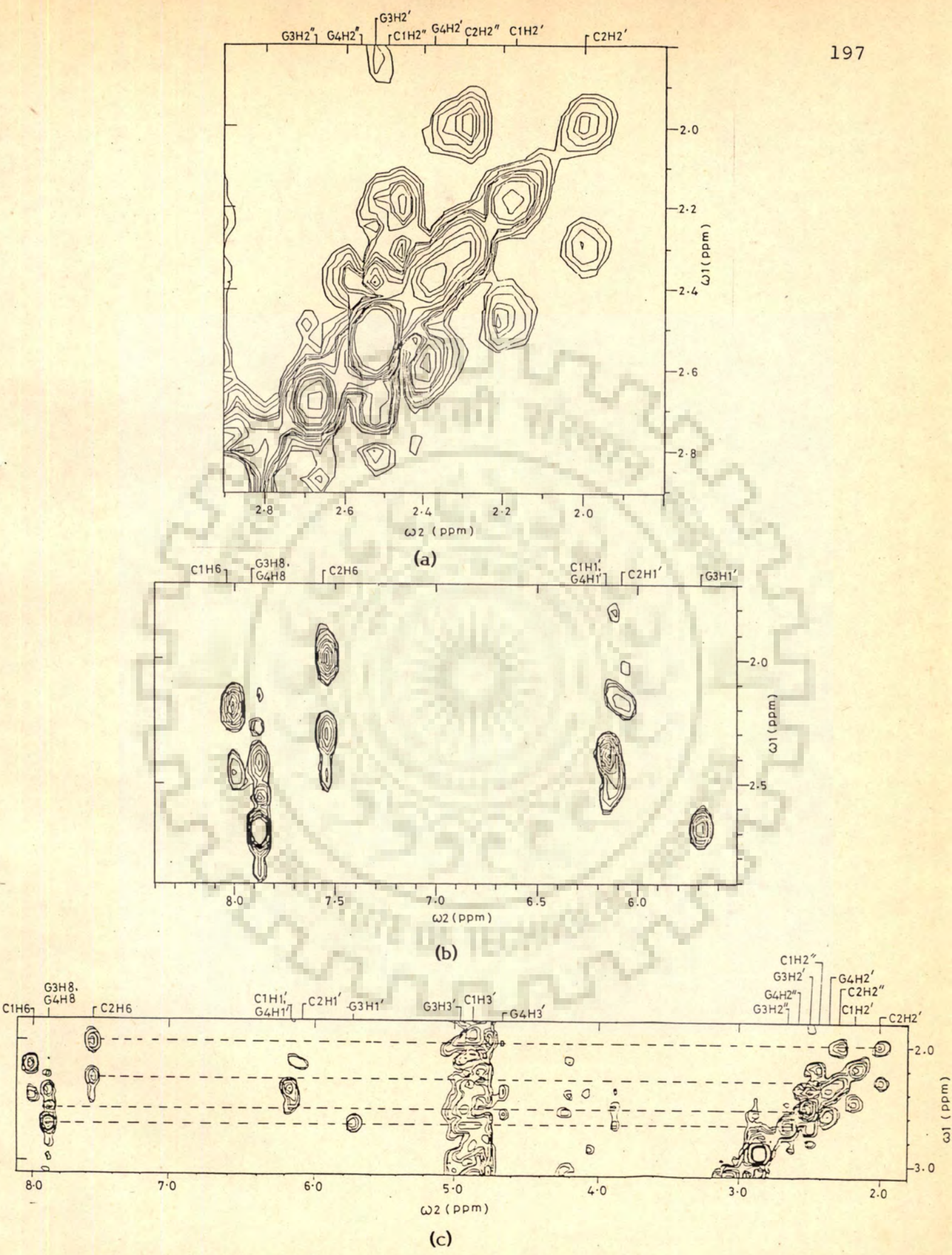


Fig. 6.4 **(a), (b)** and **(c)** Portions of NOESY spectra of fig. 6.4, shown in expanded scales

is the same as that used in Figure 6.1. It is seen that out of four triplets of H1' protons, two H1' overlap (The most downfield resonance) at 297 K. Since each one of the H1' is J-coupled to H2' and/or H2" which are further coupled to H3' and then H4' H5', H5", it is easy to find four sets of sugar resonances The various COSY and NOESY connectivities are reproduced in a more convenient form in the schematic representations of Figures 6.5 and 6.6 made by using the experimental results of Figures 6.3 and 6.4.

In order to assign a particular set of sugar protons to a specific base of d-CpCpGpG, use is made of NOE's. In a 5' - 3' sequence where n stands for nth residue, the base proton $(GH8/CH6)_n$ may show intra-residue connectivity with its corresponding H2', H2" and H1' while it is expected to show inter-residue connectivity with $(H2")_{n-1}$ and $(H1')_{n-1}$ that is, the residue preceeding it. The following are the NOE connectivities which are possible alongwith the relative intensities of cross peaks observed experimentally-

C1H6 ... C1H2'	intra-residue	strong	(1)
C1H6 ... C1H2"	- do -	weak	(2)
C1H6 ... C1H1'	- do -	Not seen	(3)
C2H6 ... C1H1'	inter-residue	weak	(4)
C2H6 ... C1H2"	- do -	weak	(5)
C2H6 ... C2H2'	intra-residue	strong	(6)

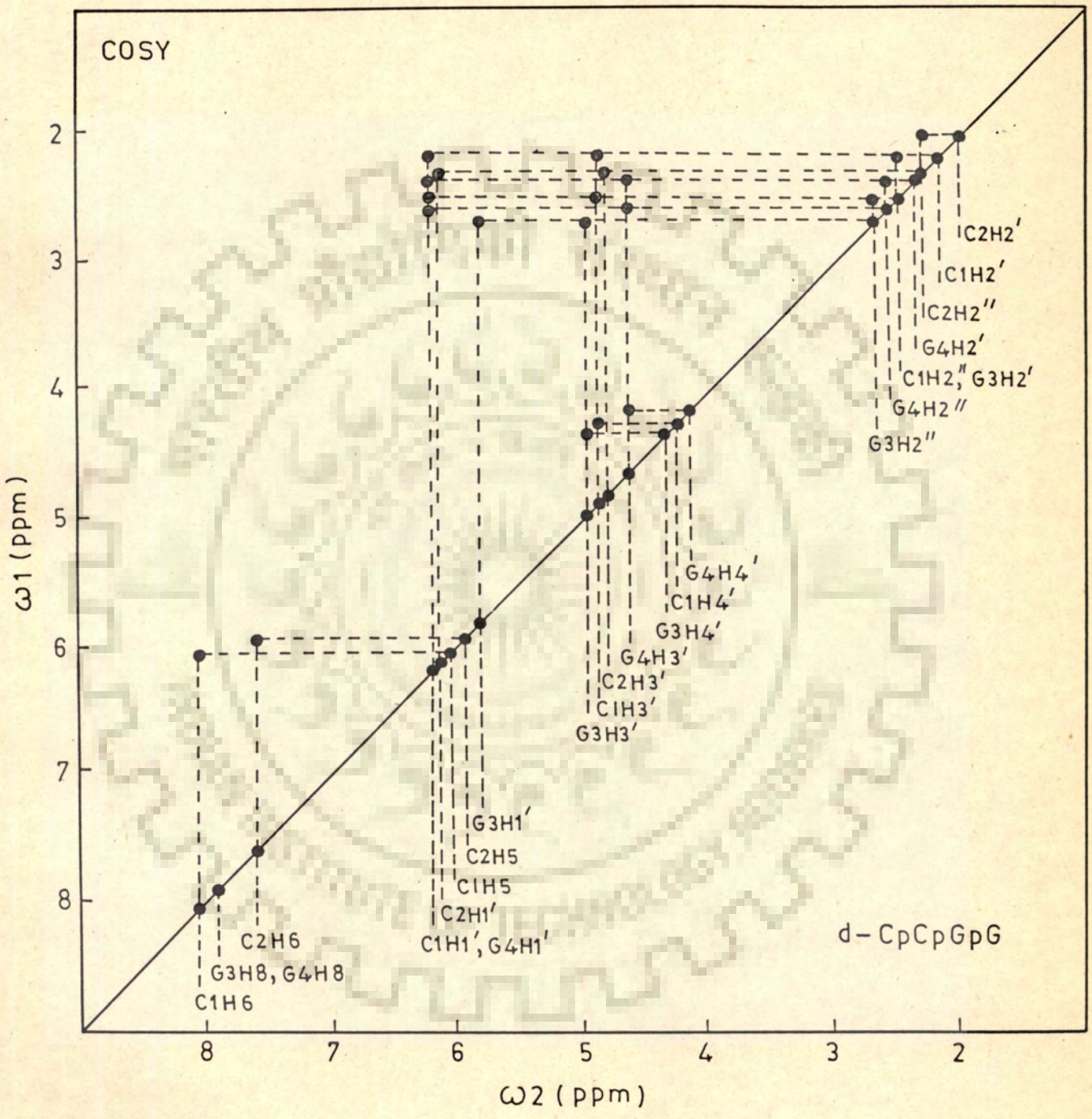


Fig. 6.5 Schematic representation of the results of COSY spectra of d-CpCpGpG shown in fig. 6.3.

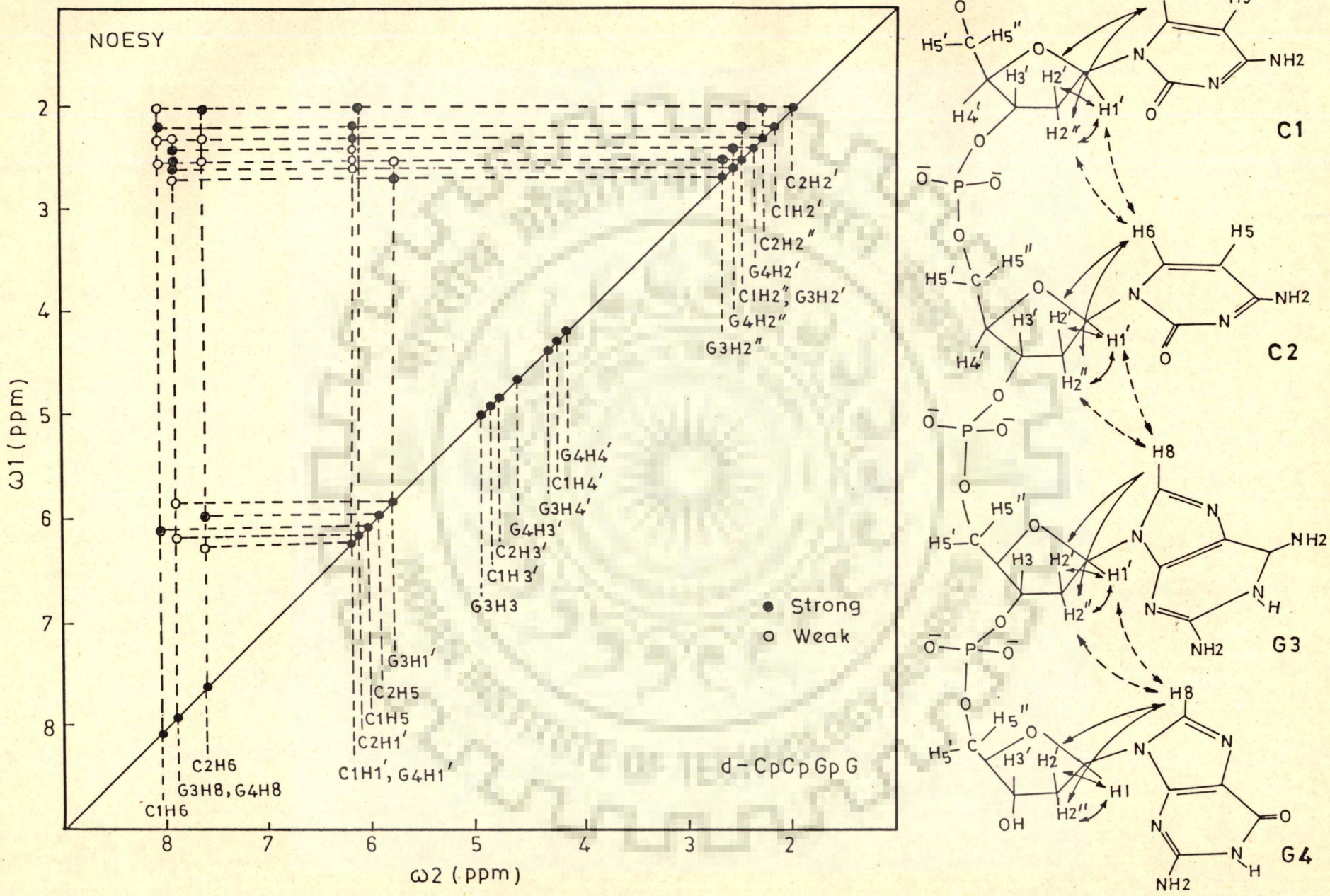


Fig. 6.6 Schematic representation of the results of NOESY spectra of d-CpCpGpG shown in fig. 6.4 alongwith 5'-3' d-CpCpGpG oligonucleotide giving various intra- and inter-residue NOE connectivities

C2H6 ... C2H2"	- do -	weak	(7)
C2H6 ... C2H1'	- do -	Not seen	(8)
G3H8 ... C2H1'	inter-residue	weak	(9)
G3H8 ... C2H2"	- do -	"	(10)
G3H8 ... G3H2'	intra-residue	strong	(11)
G3H8 ... G3H2"	- do -	weak	(12)
G3H8 ... G3H1'	- do -	Not seen	(13)
G4H8 ... G3H1'	inter -residue	weak	(14)
G4H8 ... G3H2"	- do -	weak	(15)
G4H8 ... G4H2'	intra-residue	strong	(16)
G4H8 ... G4H2"	- do -	weak	(17)
G4H8 ... G4H1'	- do -	Not seen	(18)
C1H1' ... C1H2'	- do -	strong	(19)
C1H1'... C1H2"	- do -	weak	(20)
C2H1'... C2H2'	- do -	strong	(21)
C2H1'... C2H2"	- do -	weak	(22)
G3H1'... G3H2'	- do -	weak	(23)
G3H1'... G3H2"	- do -	strong	(24)
G4H1'... G4H2'	- do -	weak	(25)
G4H1'... G4H2"	- do -	strong	(26)

The occurrence of intra-residue pairs of NOE's (1)-(2), (6)-(7), (11)-(12) and (16)-(17) assign each set of H2'-H2"

sugar protons to a particular base of d-CpCpGpG. The appearance of intra-residue connectivities (5), (10), (14) and (15) clearly assign the two CH6's to C1H6 (towards downfield out of the two CH6's) and C2H6 protons. Having done this the assignment of 4 triplets of H1' to their respective H2'-H2" is made straight away and then further to their corresponding H3', H4', H5' and H5".

B. Conformation

In order to obtain sugar conformations it is noted (see Figure 6.5) that H1' is not coupled to both H2' and H2" for all bases. However, H3' - H4' are J-coupled in all the cases. Using strategies discussed earlier each of the bases is in **O1'-endo** conformation. For bases C1 and C2, the base to H2' connectivity (19) and (21) is stronger than that of base to H2" (20) and (22) and, therefore, both have glycosidic bond rotation χ_{CN} as **anti** . However, for G3 and G4 the base to H2" (24) and (26) connectivity is stronger than base to H2' connectivity (23) and (25) so that χ_{CN} for them is in **high anti** conformation. The existence of cross peaks between base protons GH8/CH6 to their corresponding H2' [(1), (11), (6), (16)] and a set of cross peaks no. (5), (10) and (15), that is of $(GH8/CH6)_n$ to $(H2")_{n-1}$ indicates the presence of **right-handed** B-DNA structures. Also the absence of any connectivity of $(GH8)_n$ with $(CH5")_{n-1}$ shows that left-handed structure does not exist in our sample solution.

Hence it is concluded that d-CpCpGpG adopts a right handed B-DNA double helical structure at 285 K which changes to single-strand on increasing temperature with $T_{\frac{1}{2}}$ value of 310 - 312 K. The sugar conformations are predominantly 01'-endo for all bases whereas glycosidic bond rotation is anti for C1 and C2 but high anti for G3 and G4 bases.

BINDING OF d-CpCpGpG TO Lys-Tyr-Lys

Figures 6.7 and 6.8 show the results of experiments on binding of d-CpCpGpG to tripeptide Lys-Tyr-Lys at different temperatures. The changes in chemical shift on binding are given in Table 6.2. The difference between chemical shift of nucleotide protons at lowest and highest temperature has increased on adding Lys-Tyr-Lys indicating that the duplex has got more stabilised now. Consequently the $T_{\frac{1}{2}}$ value has increased as follows:

	$T_{\frac{1}{2}}$ (K) in d-CpCpGpG	$T_{\frac{1}{2}}$ (K) in d-CpCpGpG + Lys-Tyr-Lys
G4H8	310	312
C2H6	310	320
C2H1'	312	312
G3H1'	310	315
C2H5	310	310

The base protons CH6, CH5, GH8 and H1' sugar protons have all practically shifted upfield upto ~0.1 ppm at

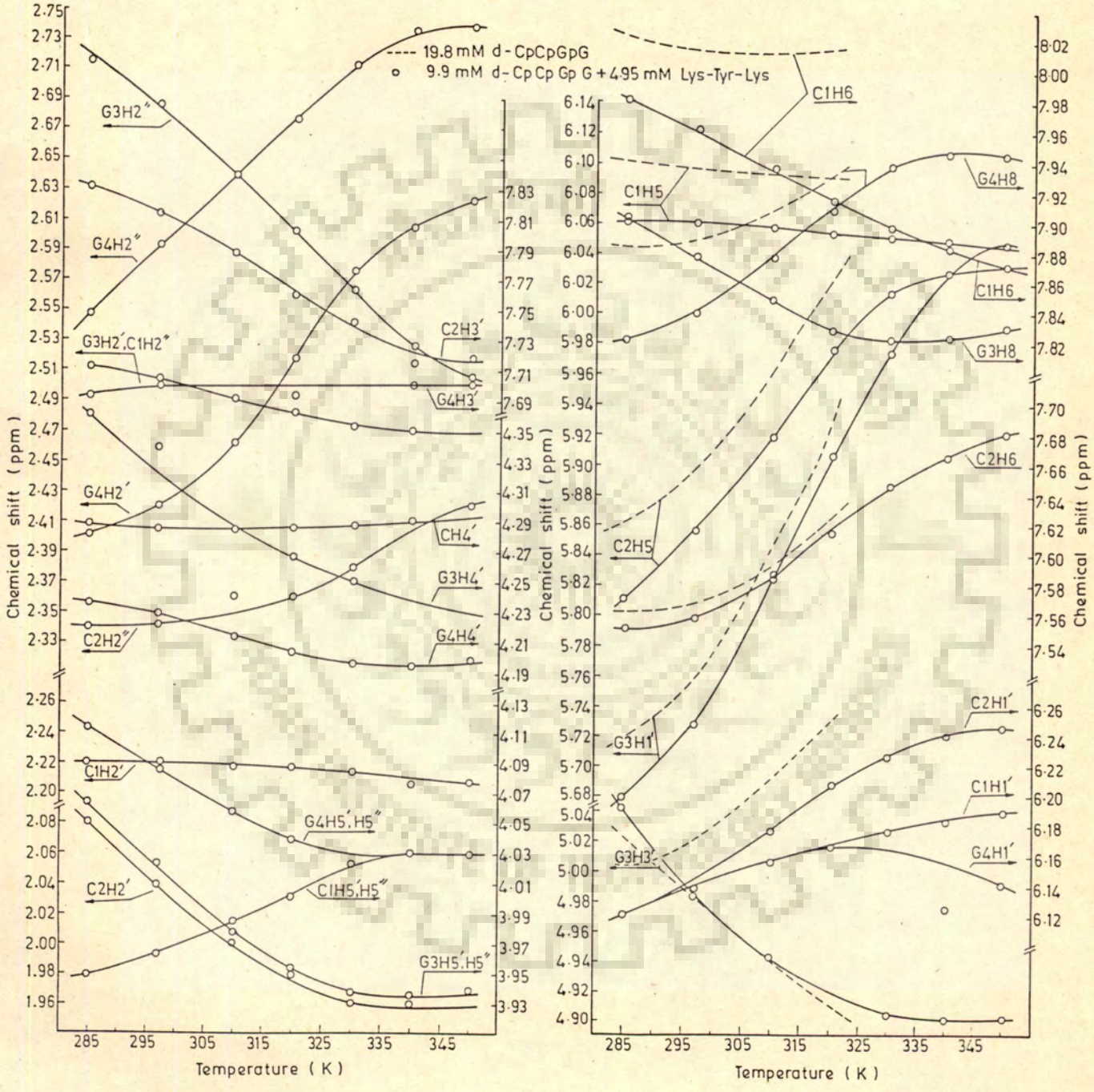


Fig. 6.8 Chemical shifts of some of the nucleotide and peptide protons as a function of temperature in uncomplexed nucleotide/peptide and the complexed state from the data in fig. 6.7.

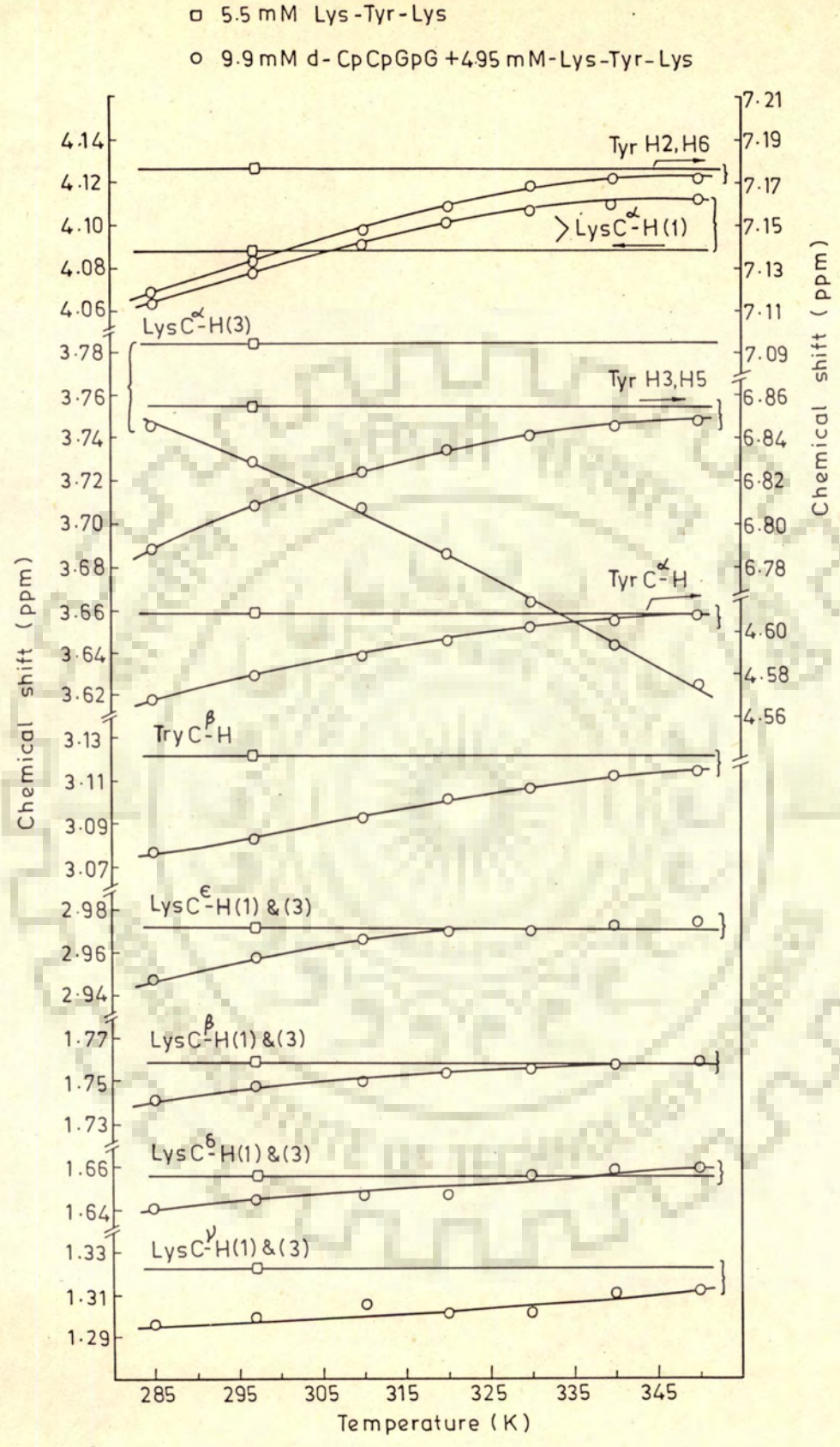

□ 5.5 mM Lys-Tyr-Lys
○ 9.9 mM d-CpCpGpG +4.95 mM-Lys-Tyr-Lys
Tyr H2,H6
Lys Cα-H(1)
Lys Cα-H(3)
Tyr H3,H5
Tyr Cα-H
Try Cβ-H
Lys Cε-H(1) & (3)
Lys Cβ-H(1) & (3)
Lys Cδ-H(1) & (3)
Lys Cγ-H(1) & (3)
Chemical shift (ppm)
Chemical shift (ppm)
Temperature (K)
4.14
4.12
4.10
4.08
4.06
3.78
3.76
3.74
3.72
3.70
3.68
3.66
3.64
3.62
3.13
3.11
3.09
3.07
2.98
2.96
2.94
1.77
1.75
1.73
1.66
1.64
1.33
1.31
1.29
7.21
7.19
7.17
7.15
7.13
7.11
7.09
6.86
6.84
6.82
6.80
6.78
4.60
4.58
4.56
285
295
305
315
325
335
345

Fig. 6.8 Contd.

Table 6.2 : Changes in chemical shift $\Delta\delta$ in ppm of peptide and nucleotide protons on binding of 9.9 mM d-CpCpGpG to 4.95 mM Lys-Tyr-Lys at indicated temperatures

SYSTEM	Temp. (K)	G3H8	G4H8	C1H6	C2H6	C1H5	C2H5	C1H1'	C2H1'	G3H1'
	285	7.926	7.889	8.026	7.575	6.102	5.863	6.218	6.156	5.720
d-CpCpGpG	297	7.887	7.887	8.013	7.566	6.097	5.896	6.206	6.168	5.755
	320	7.824	7.932	8.012	7.627	6.090	6.020	6.188	6.249	5.925
	285	7.903	7.821	7.982	7.552	6.060	5.811	6.121	6.121	5.677
d-CpCpGpG + Lys-Tyr-Lys	297	7.879	7.839	7.963	7.558	6.061	5.856	6.138	6.138	5.726
	320	7.827	7.909	7.915	7.615	6.053	5.977	6.166	6.208	5.906
	285	0.023	0.068	0.044	0.023	0.042	0.052	0.097	0.035	0.043
$\Delta\delta$ (ppm)*	297	0.008	0.048	0.050	0.008	0.036	0.040	0.068	0.030	0.029
	320	-0.003	0.023	0.097	0.012	0.037	0.043	0.022	0.041	0.019

SYSTEM	Temp. (K)	C1H2'	$\frac{C1H2"}{G3H2'}$	G3H2"	C2H2'	C2H2"	G4H2'	G4H2"	C1H3'	C2H3'	G3H3'
	285	2.242	2.522	2.745	2.100	2.359	2.406	2.554	-	-	5.025
d-CpCpGpG	297	-	2.494	2.692	2.047	2.342	2.412	2.597	-	-	4.986
	320	2.226	2.446	2.604	1.945	2.343	2.478	2.690	4.799	4.756	-
	285	2.220	2.514	2.714	2.080	2.340	2.400	2.546	-	-	5.045
d-CpCpGpG + Lys-Tyr-Lys	297	-	2.503	2.686	2.040	2.340	2.420	2.593	-	-	4.932
	320	2.219	2.477	2.604	1.980	2.360	2.517	2.676	4.797	4.758	-
	285	0.022	0.008	0.031	0.020	0.019	0.006	0.008	-	-	-0.020
$\Delta\delta$(ppm)*	297	-	-0.009	0.006	0.007	0.002	-0.008	0.004	-	-	0.004
	320	0.007	-0.031	0.000	-0.035	-0.017	-0.039	0.014	0.002	-0.002	-

Contd.

SYSTEM	Temp. (K)	G4H3'	C1H4'	G3H4'	G4H4'	C1H5'/H5"	G3H5'/H5"	G4H5'/H5"
d-CpCpGpG	285	4.702	4.279	4.375	4.263	3.918	4.040	4.116
	297	4.713	4.277	4.344	4.234	3.918	4.024	4.103
	320	4.728	4.288	4.262	4.177	3.928	4.025	4.091
d-CpCpGpG + Lys-Tyr-Lys	285	4.691	4.288	4.360	4.235	3.948	4.063	4.112
	297	4.700	4.285	4.338	4.228	3.963	4.024	4.085
	320	4.693	4.287	4.266	4.203	4.004	3.954	4.038
Δδ (ppm)*	285	0.011	-0.009	0.015	0.028	-0.030	-0.023	0.004
	297	0.013	-0.008	0.006	0.006	-0.045	0.000	0.018
	320	0.027	0.001	-0.001	-0.026	-0.076	0.071	0.053

SYSTEM	Temp. (K)	Tyr H2,H6	Tyr H3,H5	Tyr C^{α}-H	Tyr C^{β}-H	Lys $^{\alpha}C$-H (1)	Lys $^{\alpha}C$-H (3)	Lys $^{\beta}C$-H (1)&(3)	Lys $^{\gamma}C$-H (1)&(3)	Lys $^{\delta}C$-H (1)&(3)	Lys $^{\varepsilon}C$-H (1)&(3)
Lys-Tyr-Lys	297	7.177	6.854	4.608	3.121	4.089	3.785	1.758	1.323	1.655	2.971
	285	7.117	6.789	4.568	3.077	4.063	3.745	1.741	1.296	1.641	2.947
d-CpCpGpG + Lys-Tyr-Lys	297	7.133	6.808	4.580	3.083	4.078	3.729	1.747	1.299	1.644	2.958
	320	7.155	6.833	4.592	3.102	4.100	3.685	1.753	1.301	1.646	2.969
Δδ(ppm)*	285	0.060	0.065	0.040	0.044	0.026	0.040	0.017	0.027	0.014	0.024
	297	0.044	0.046	0.028	0.038	0.011	0.056	0.011	0.024	0.011	0.013
	320	0.022	0.021	0.016	0.019	-0.011	0.100	0.005	0.022	0.009	0.002

*Upfield shifts, Δδ in ppm, are taken as positive sign.

285 K. Tyr ring protons shift upfield by about 0.06 ppm at 285 K. Further, it may be noted that the change in chemical shift in all cases decreases with temperature. This once again demonstrates that binding preferentially occurs to the double-stranded d-CpCpGpG than that to the single-strand. The upfield shift in Tyr ring protons is due to stacking/intercalation/partial stacking on/between base pairs of d-CpCpGpG. However, it is to a lesser extent ($\Delta\delta \sim 0.06$ ppm) in d-CpCpGpG than that to d-GpCpGpC ($\Delta\delta \sim 0.14$ ppm).

Figures 6.9 and 6.10 show results of 2D-COSY and NOESY spectra of a mixture of d-CpCpGpG with Lys-Tyr-Lys for the same sample solution for which 1D-NMR have been shown in Figure 6.7 - 6.8. Figures 6.9a, b, c and 6.10a, b are the portions of these spectra on expanded scales. Figures 6.11 and 6.12 are schematic representations of these results. For all the bases the connectivities within various sugar resonances is found to be same as that in uncomplexed state. The sugar conformation is, therefore, predominantly 01'-endo.

The relative intensities of NOE cross peaks in d-CpCpGpG bound to Lys-Tyr-Lys are as follows:

C1H6 ..... C1H2'	intra-residue	strong	(31)
C1H6 ..... C1H2"	"	nil	(32)
C1H6 ..... C1H1'	"	nil	(33)

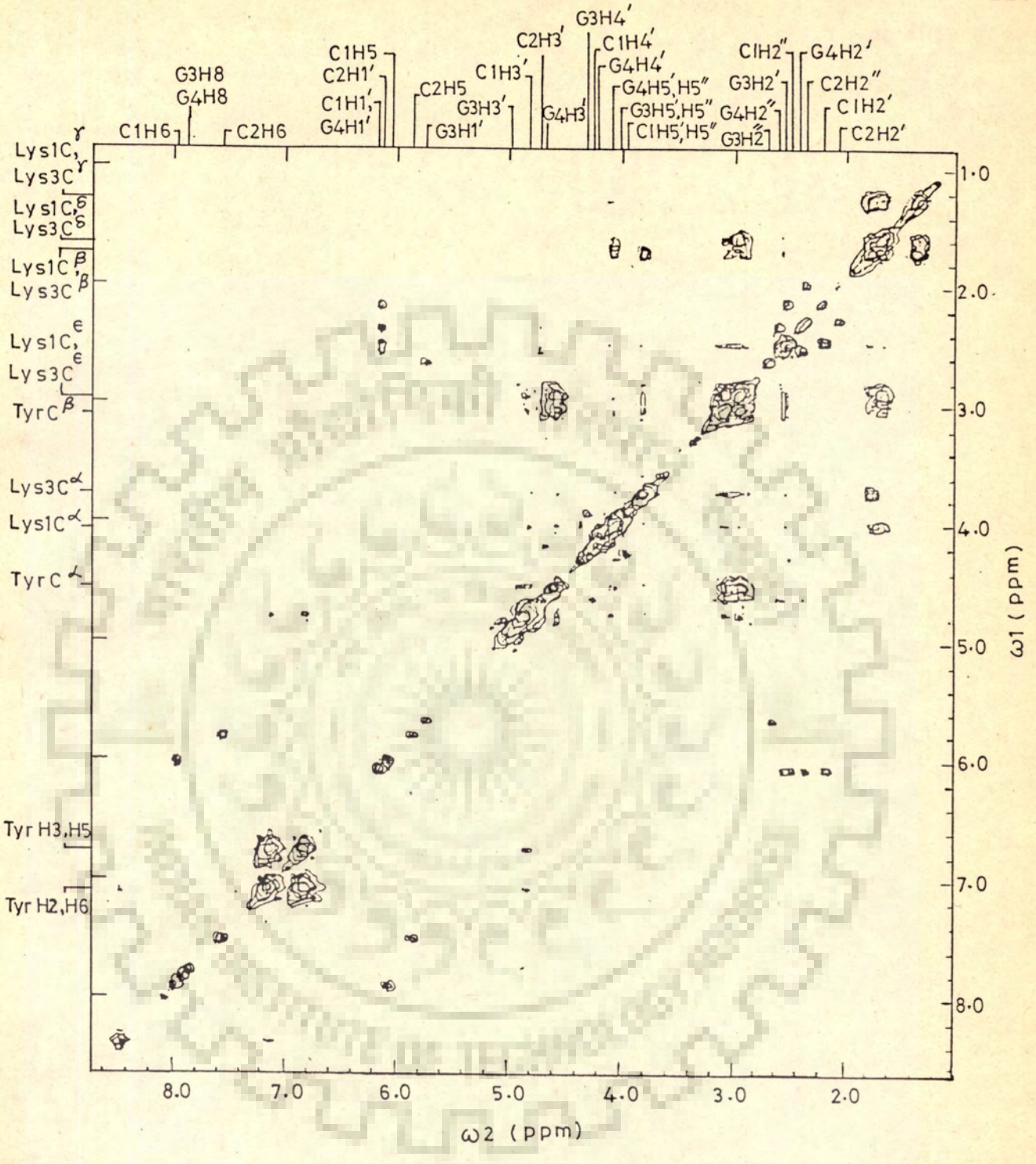


Fig. 6.9 500 MHz COSY spectrum of a mixture of 9.9 mM d-CpCpGpG and 4.95 mM Lys-Tyr-Lys in D_2O at 297 K for the sample used in fig.6.7.

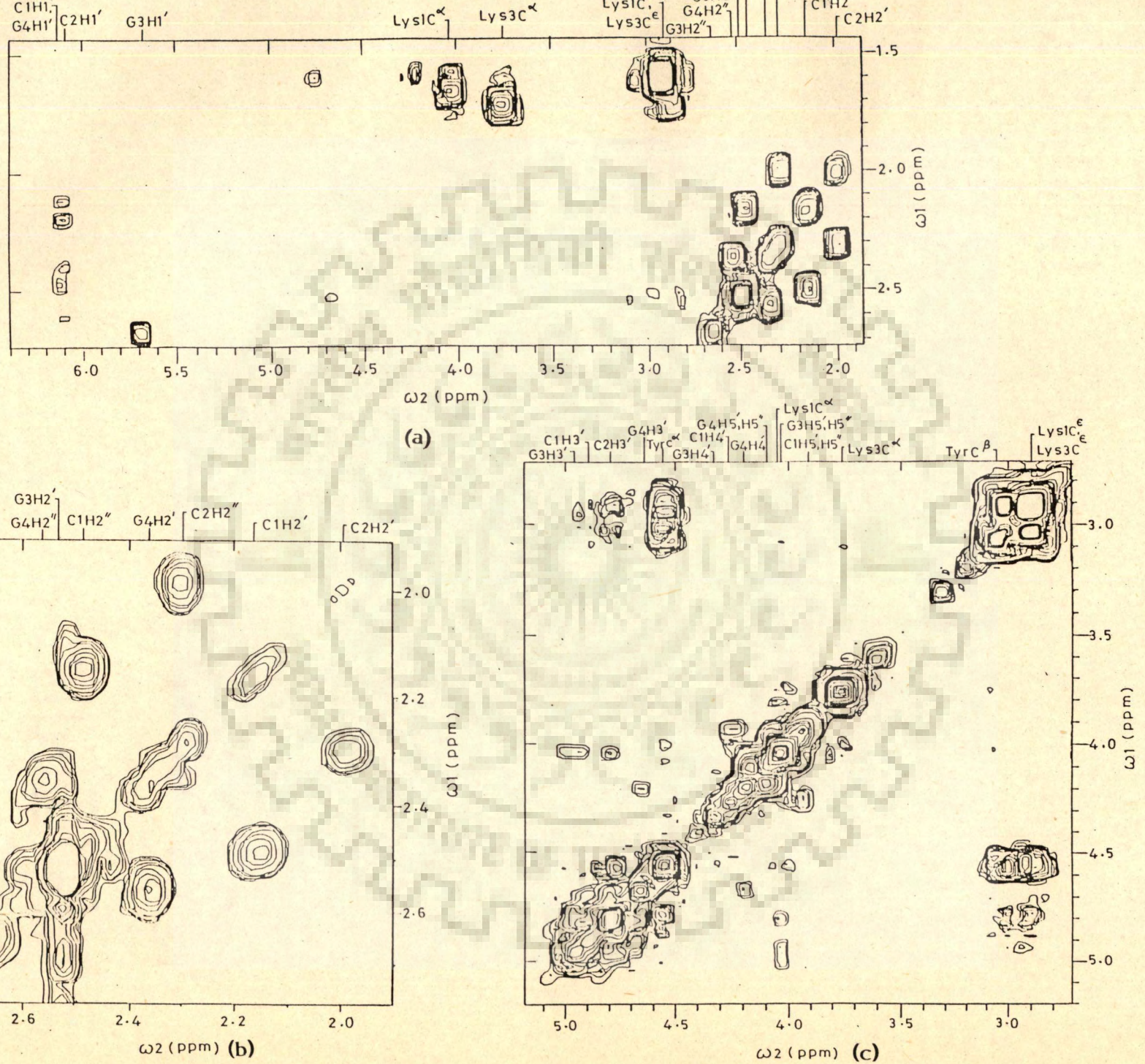


Fig. 6.9 (a), (b), (c) Portions of COSY spectra of fig. 6.9 in an expanded scale

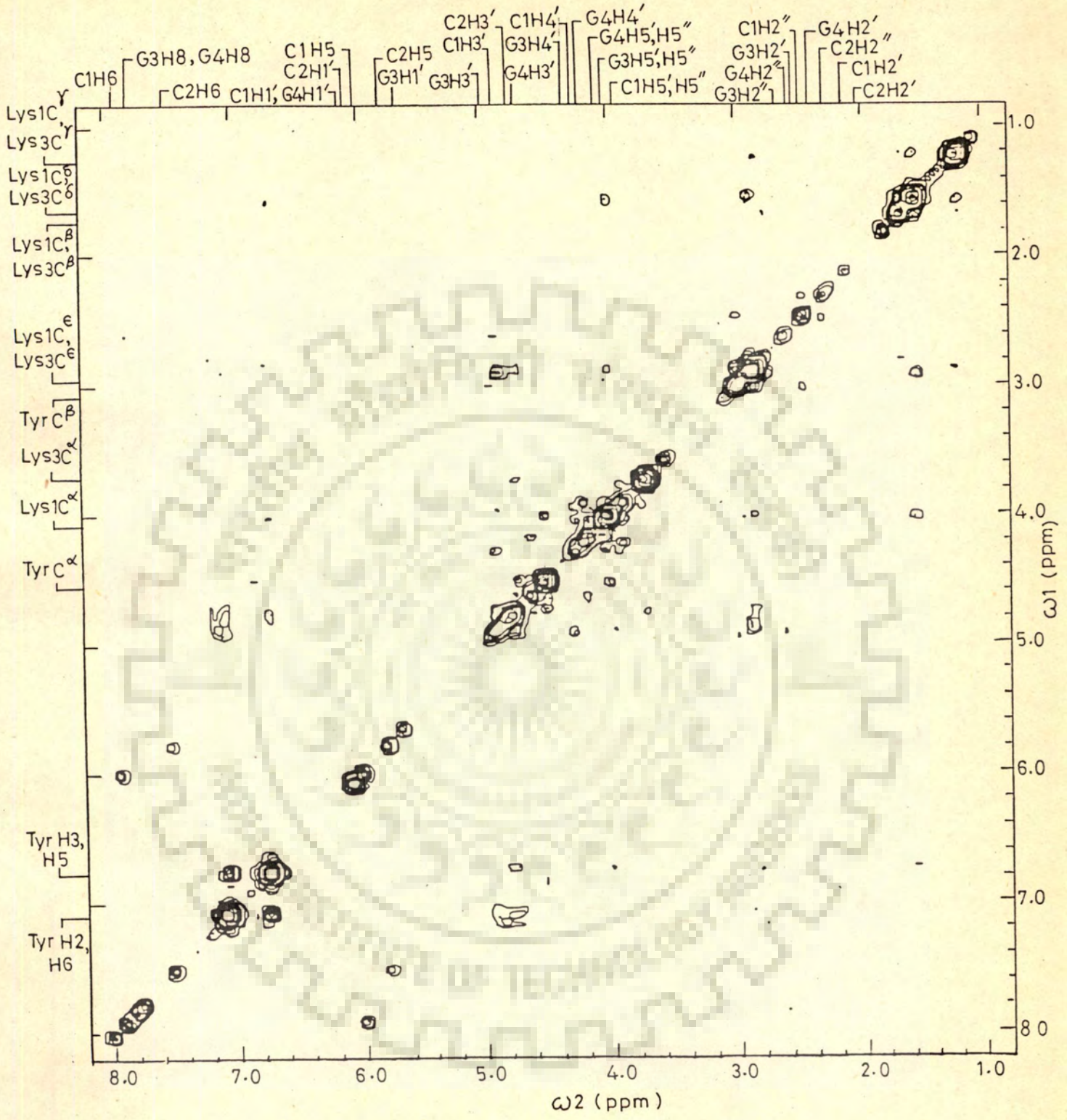
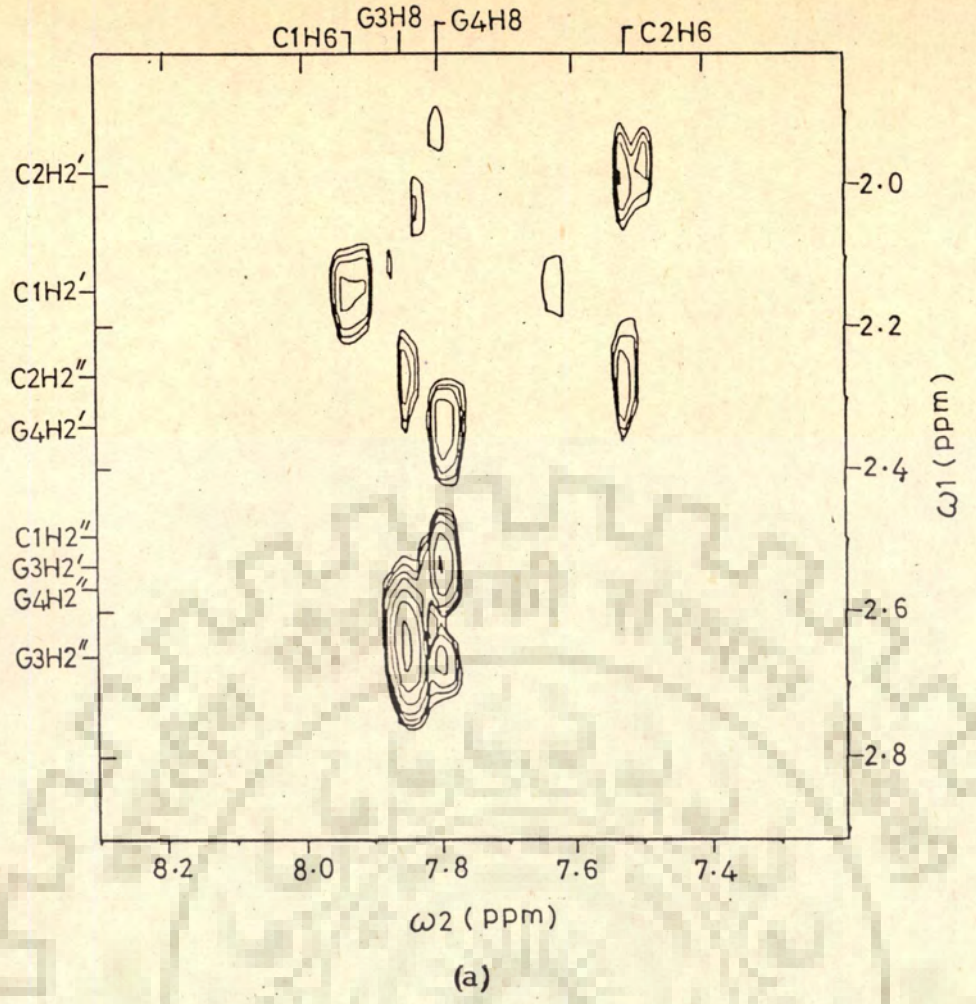


Fig. 6.10 500 MHz NOESY spectrum of a mixture of 9.9 mM d-CpCpGpG and 4.95 mM Lys-Tyr-Lys in D_2O at 297 K for the sample used in fig.6.7.



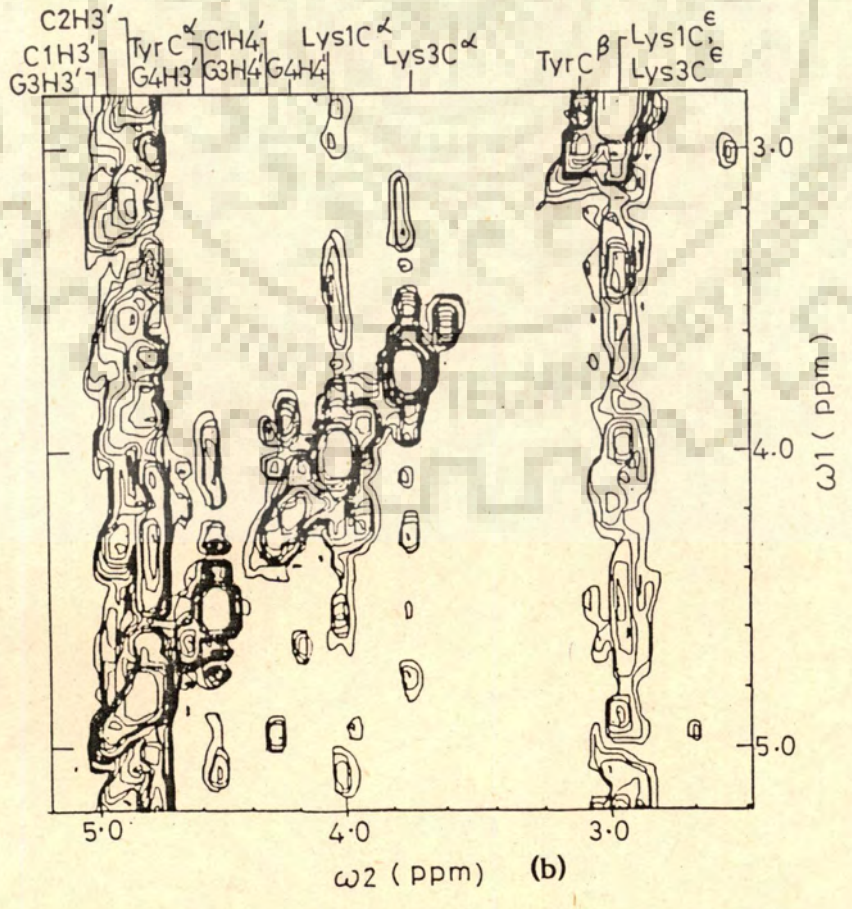


Fig. 6.10 (a) (b) Portions of NOESY spectrum of Fig.6.10 in an expanded scale.

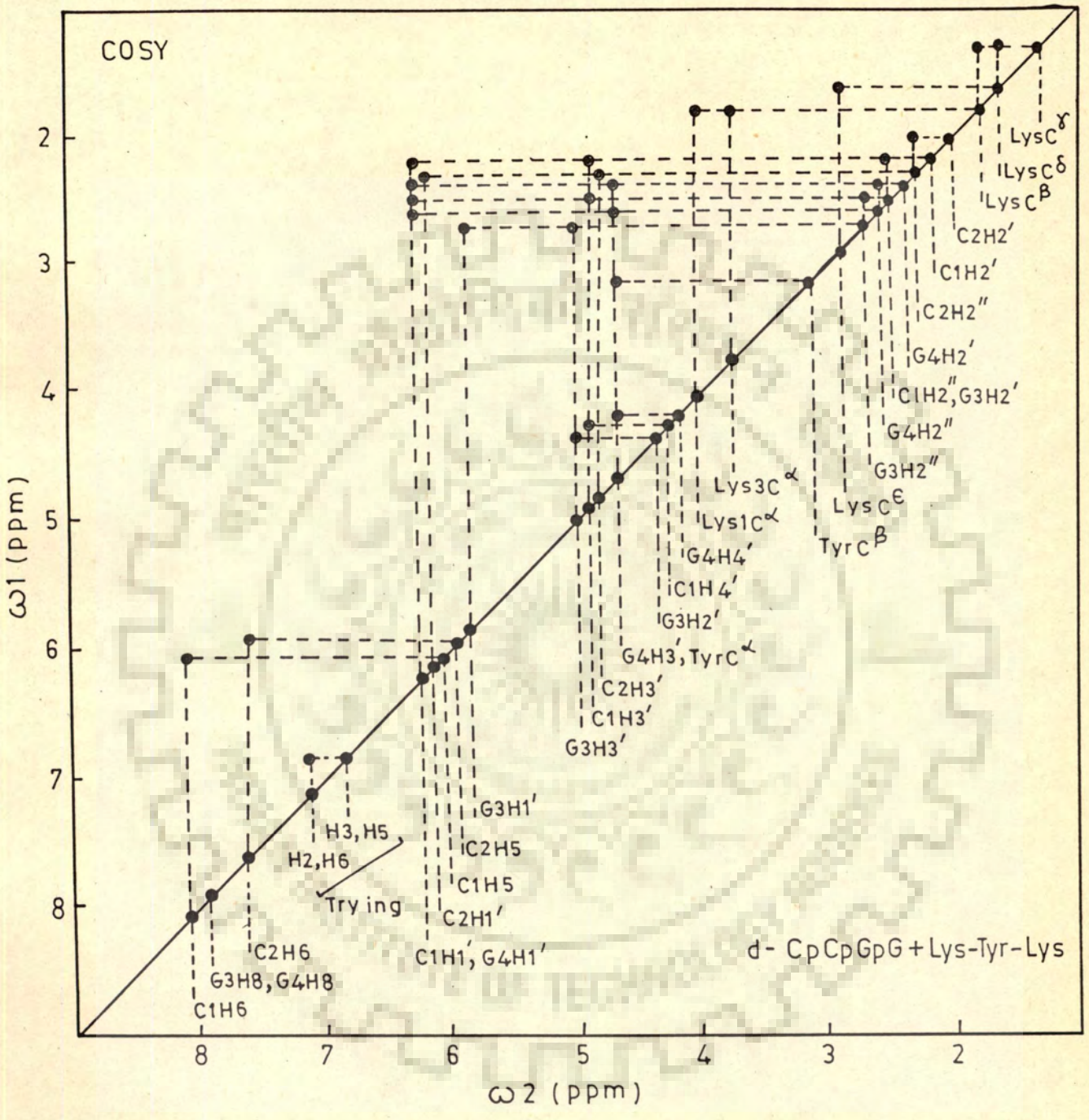


Fig. 6.11 Schematic representation of the results of COSY spectra of mixture of d-CpCpGpG and Lys-Tyr-Lys taken from fig. 6.9.

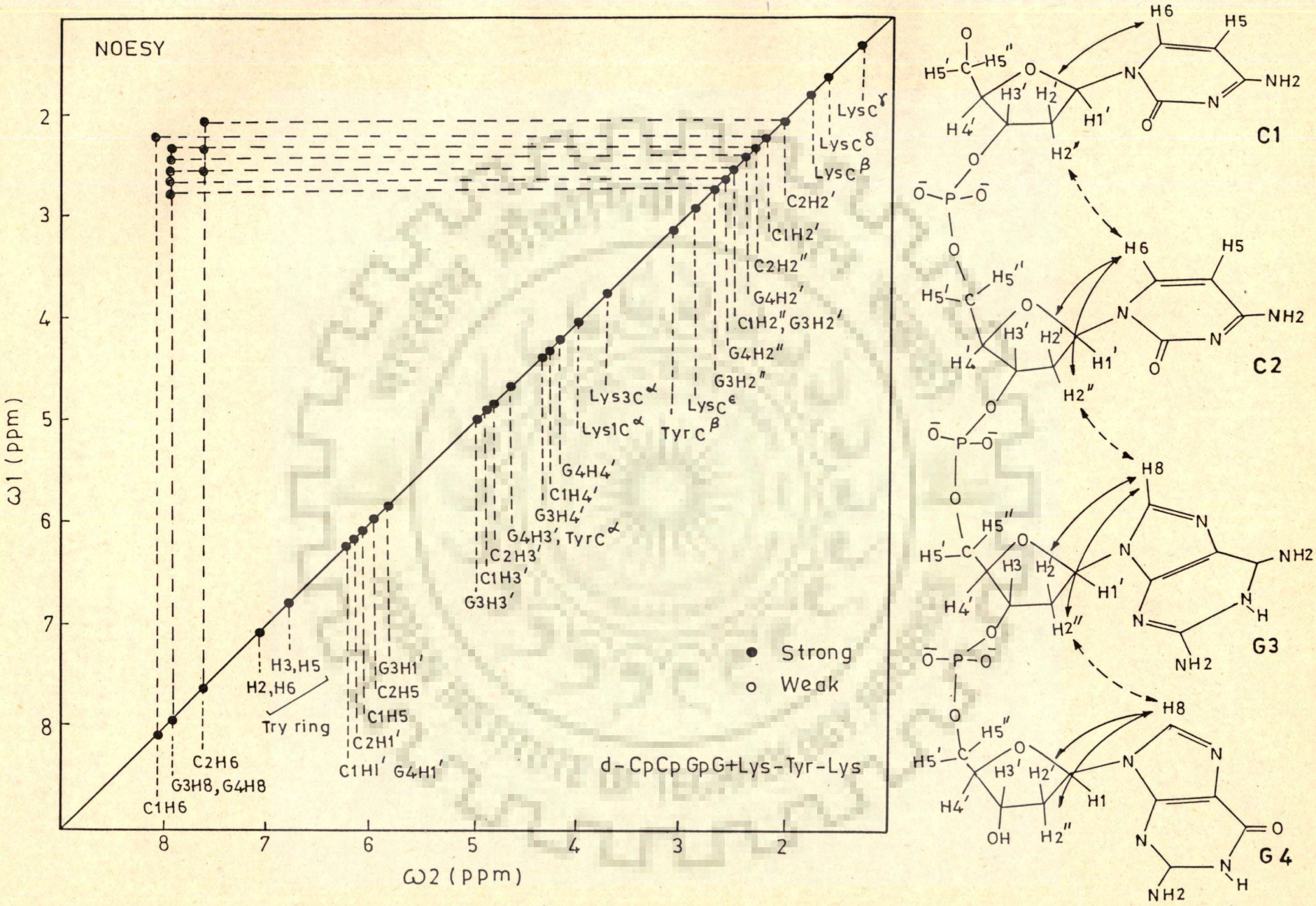


Fig. 6.12 Schematic representation of the results of NOESY spectra of mixture d-CpCpGpG and Lys-Tyr-Lys taken from fig. 6.10.

C2H6	....	C1H1'	inter-residue	nil	(34)
C2H6	....	C1H2"	"	strong	(35)
C2H6	....	C2H2'	intra-residue	strong	(36)
C2H6	....	C2H2"	"	strong	(37)
C2H6	....	C2H1'	"	nil	(38)
G3H8	....	C2H1'	inter-residue	nil	(39)
G3H8	....	C2H2"	"	strong	(40)
G3H8	....	G3H2'	intra-residue	strong	(41)
G3H8	....	G3H2"	"	strong	(42)
G3H8	....	G3H1'	"	nil	(43)
G4H8	....	G3H1'	inter-residue	nil	(44)
G4H8	....	G3H2"	"	strong	(45)
G4H8	....	G4H2'	intra-residue	strong	(46)
G4H8	....	G4H2"	"	strong	(47)
G4H8	....	G4H1'	"	nil	(48)
C1H1'	....	C1H2'	"	nil	(49)
C1H1'	....	C1H2"	"	nil	(50)
C2H1'	....	C2H2'	"	"	(51)
C2H1'	....	C2H2"	"	"	(52)
G3H1'	....	G3H2'	"	"	(53)
G3H1'	....	G3H2"	"	"	(54)
G4H1'	....	G4H2'	"	"	(55)
G4H1'	....	G4H2"	"	"	(56)

For all the four bases the base proton to corresponding H2' proton has stronger NOE than to corresponding H2" protons. This is clear from pairs of peak Nos. (31)-(32), (36)-(37), (41)-(42), and (46)-(47) for bases C1, C2 and G3 and G4 respectively. This indicates that χ_{CN} is in anti conformation.

Further inter-residue NOE'S (35), (40), (45) further support the argument that χ_{CN} is anti in all cases. There are no clear inter-strand NOE'S which could demonstrate that complex with d-CpCpGpG is in duplex state. Intra-residue NOE'S of base protons with corresponding H2' protons i.e. peak No.(31), (36), (41) and (46) inter-residue NOE'S No. (35), (40) and (45) support the hypothesis that the DNA in complexed form is right-handed. There are no intermolecular NOE'S between peptide protons and nucleic acid protons which could give direct evidence that two are close to each other in space or form a stacked complex.

The conformation of the complex of d-CpCpGpG with Lys-Tyr-Lys thus has all sugars with O1'-endo pucker, all χ_{CN} to be anti and helix sense as right-handed.

CHAPTER - 7

STACKING ENERGY CALCULATIONS FOR TRP, TYR AND PHE

CHARGE DISTRIBUTION ON INTERACTING MOLECULES

The structures of nucleic acid bases and base pairs have been taken from the results of single crystal X-Ray diffraction studies [143]. The geometries of Trp, Tyr, Phe have been taken from the published data [107, 41, 55, 36]. The charge distributions of these molecules have been calculated by CNDO (Complete neglect of differential overlap) method [114, 115]. The partial charges on the various atoms of these molecules obtained are shown in Fig. 7.1.

MINIMUM ENERGY CONFORMATION

In order to find the minimum energy and hence most stable geometry of two stacked molecules, the relative orientation of molecules, initially was as shown in Fig. 7.2(a) and the following algorithm was used.In the orthogonal coordinate system XYZ, the base moiety/base pair was held, rigid such that the planar base ring was in the X-Y plane and the Z—axis passes approximately through the centre of base moiety. Initially aromatic amino acid is oriented parallel to the X-Y plane at a separation of 3.4 Å along the Z—axis. Intermolecular interaction energy is calculated as a sum of electrostatic, polarisation, dispersion and repulsion energies as detailed in Chapter—2. The configura-

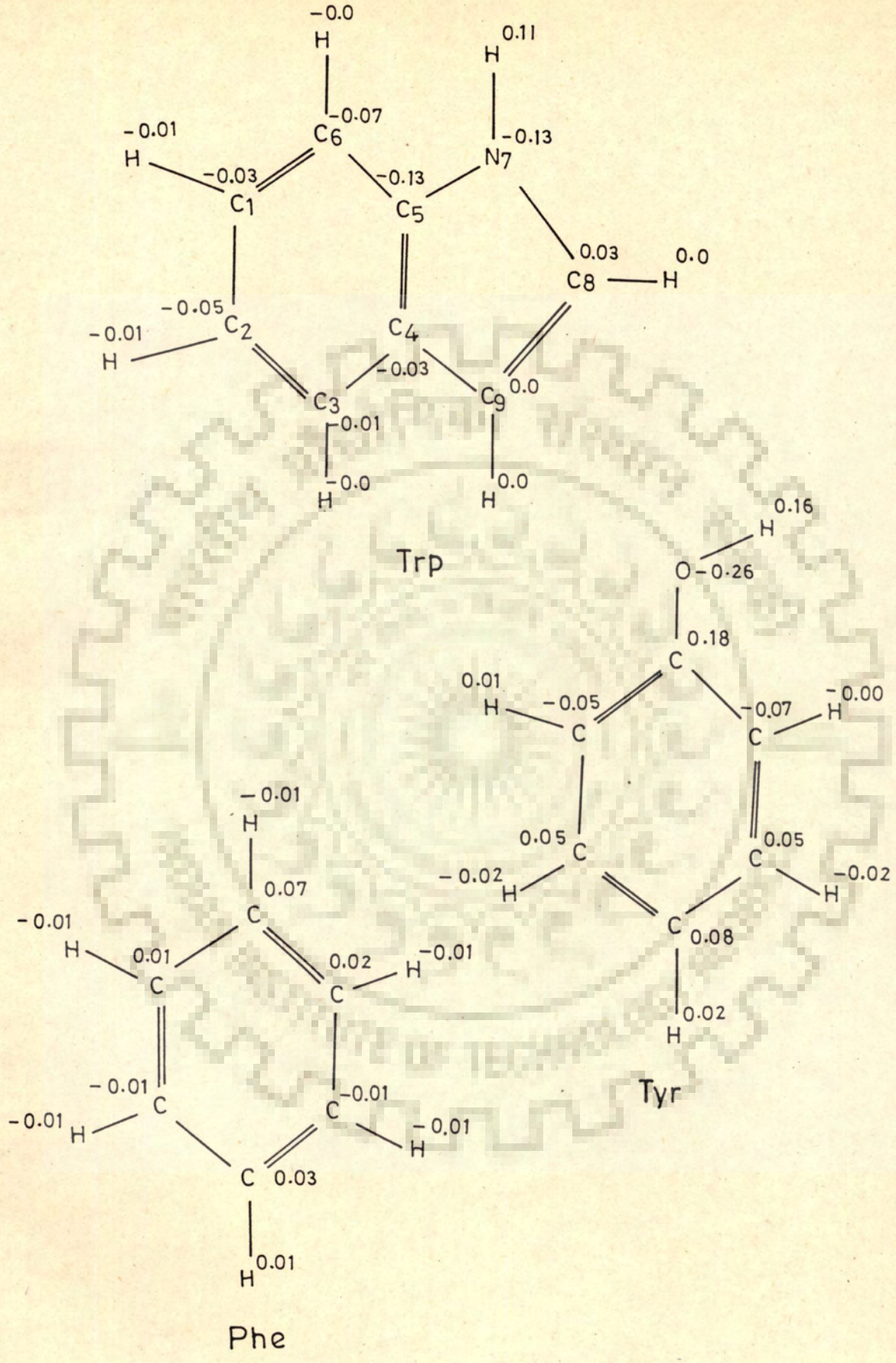


Fig. 7.1 Charge distribution on Trp, Tyr, Phe; bases C, G; base pairs C-G and A-T

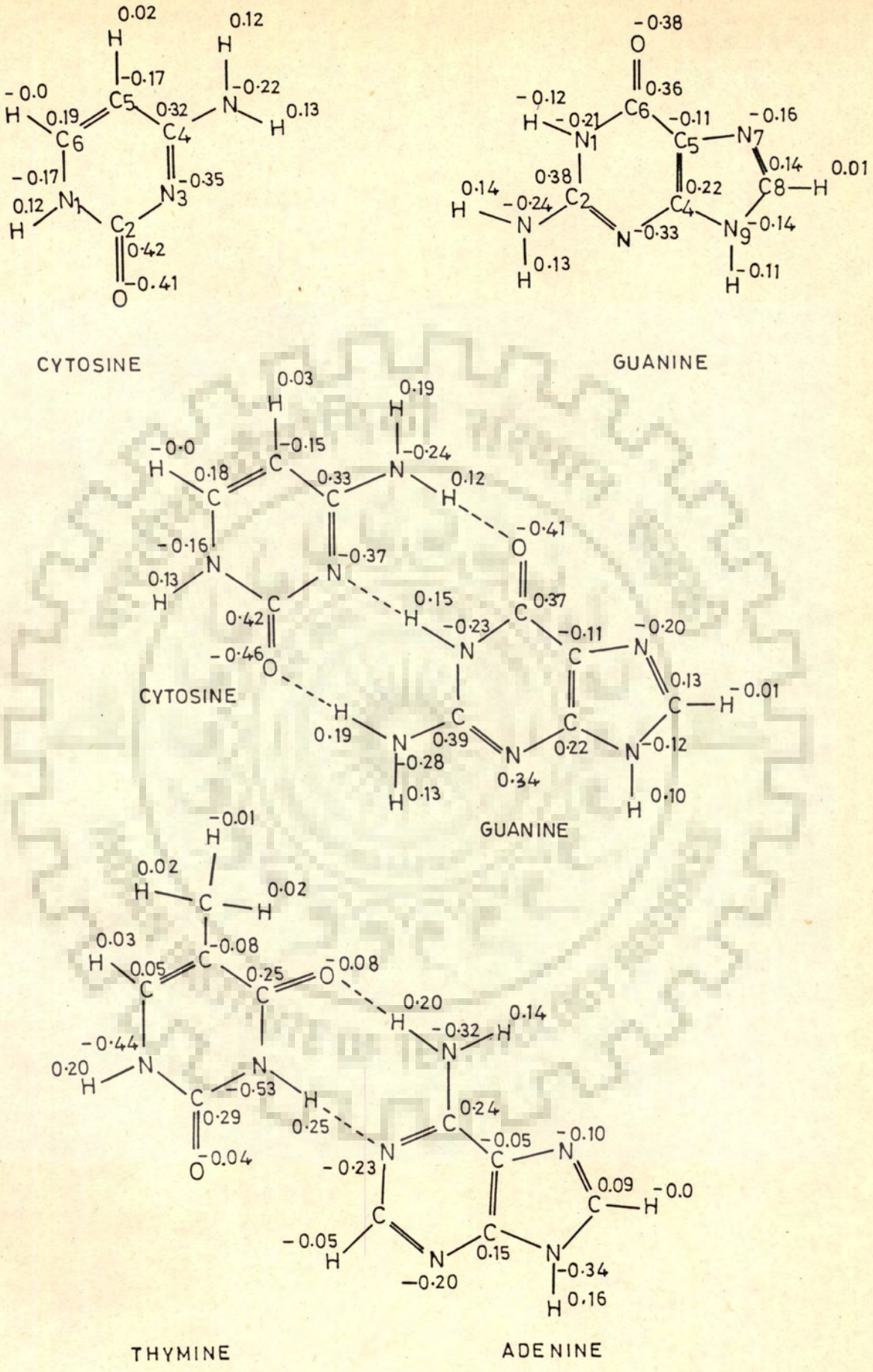


Fig. 7.1 Contd.

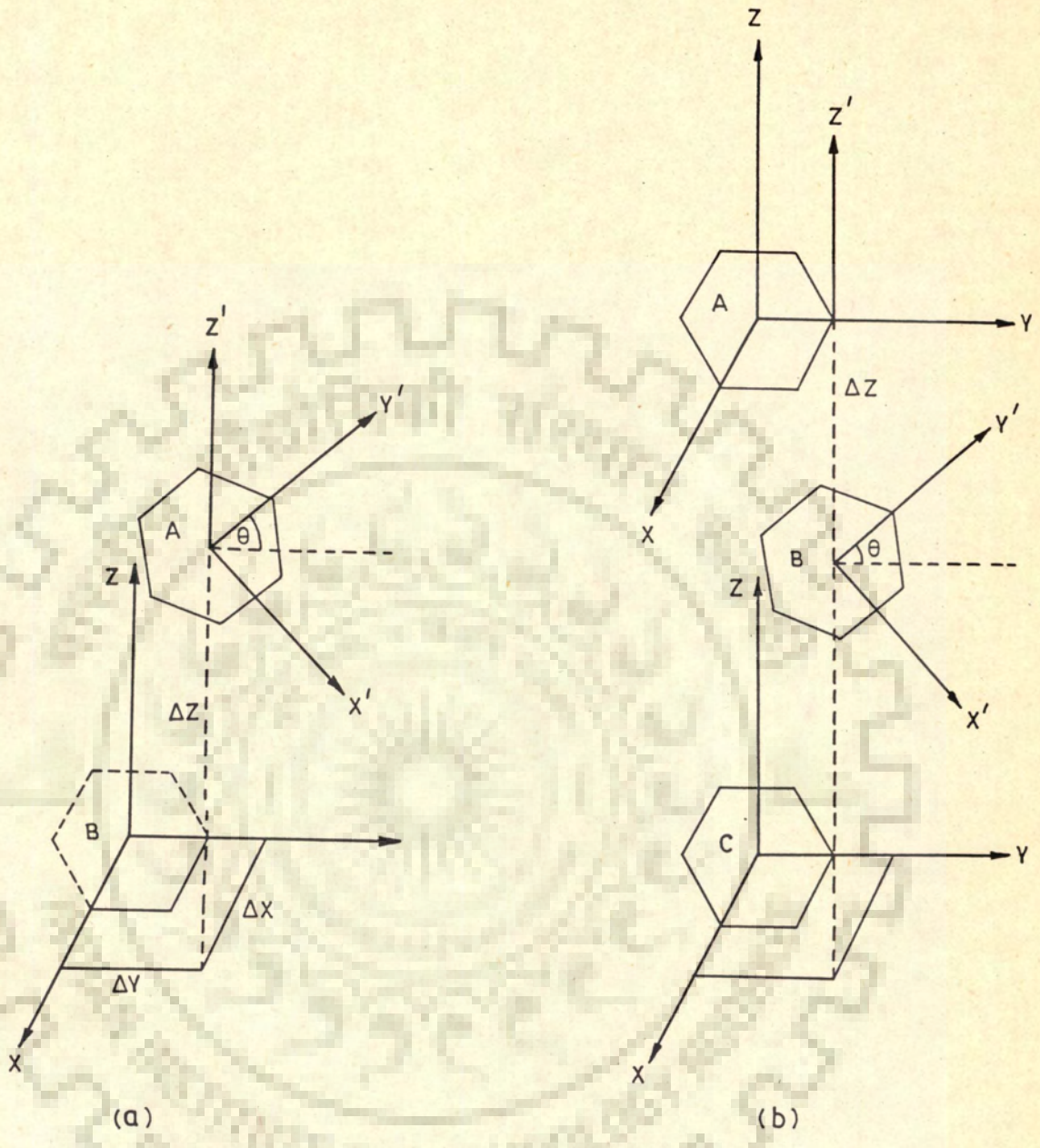


Fig. 7.2 Schematic representation showing relative orientation of (a) base (B) with respect to aromatic amino acid (A) and (b) two base pairs with aromatic amino acid intercalated between them

tional space is then scanned by rotating the aromatic amino acid about Z—axis at an interval of 10 degrees ($\Delta\theta$) and sliding it along X and Y—axis for each orientation by a displacement of 0.1 Å (X, Y). The minimum energy configuration so obtained is further optimised by varying the separation of the molecules along Z—axis by displacements at intervals of 0.1 Å.

To find interaction energy of a model system in which aromatic moiety of peptide is intercalated in duplex of d-CpG, the geometrical configuration shown in Fig. 7.2(b) was used. The aromatic ring of Trp, Tyr and Phe was placed exactly in the middle of C-G and G-C base pairs and the distance between base pairs was varied from 6.2 to 7.0 Å. The search for minimum energy configuration was then made by varying $\Delta\theta$, ΔX, ΔY and ΔZ. All energy calculations were made for two rotational angles (α) between base pairs, that is, 36° and 61°. A relative angle of 61° between two base pairs was selected since an unwinding of helix [135] is known to occur when the vertical separation between base pairs is increased from 3.4 Å (native DNA) to 6.8 Å due to intercalation of aromatic ring of drug molecules. For actinomycin, EthBr, proflavin and related molecules [133, 141, 120, 124] $\Delta\alpha$ was found to be -21° to -29°. Since no such crystallographic data is available in literature for peptides, an average value of $\Delta\alpha$ = - 25° and hence rotation of 36° + 25° = 61° was selected as a representative one

for present studies.

INTERACTION WITH BASES

The stacking energies for the minimum energy conformation of the model complexes of base/base pair/dinucleotide d-CpG with Trp, Tyr and Phe are summarized in Table 7.1. The contributions of various terms to the total interaction energy are given in Table 7.2-7.4. The only set of theoretical results available in literature, that is, those of Govil and coworkers [78] are also shown in these tables for comparison.

In all calculations the minimum energy was obtained when amino acid residue is at a distance of 3.4 Å from base. The stacking energies for bases G and C are found to be in the range -5.4 to -9.5 kcal/mole which is comparable to that found for hydrogen bonding with bases i.e. 6-15 kcal/mole [77] and lesser than those for electrostatic interaction i.e. 23-38 kcal/mole [76] and coulombic interaction [76] being about 90-130 kcal/mole. Similar values of stacking energies have earlier been reported in literature [78]. The partitioning of stacking energies show that in all cases, the contributions from dispersion terms are the dominating ones and are significantly greater than electrostatic, polarisation and repulsion energies. This is characteristic of stacking interactions. The energy of interaction

Table 7.1 : Summary of results obtained on stacking energies

SYSTEM	Trp	Tyr	Phe
C	-9.46	-8.83	-5.38
G	-7.99	-9.98	-7.95
G-C	-10.59	-8.94	-7.20
A-T	-8.63	-9.10	-7.44
d-CG (α=36°)	-20.77	-18.77	-14.95
d-CG (α=61°)	-20.16	-18.21	-14.72

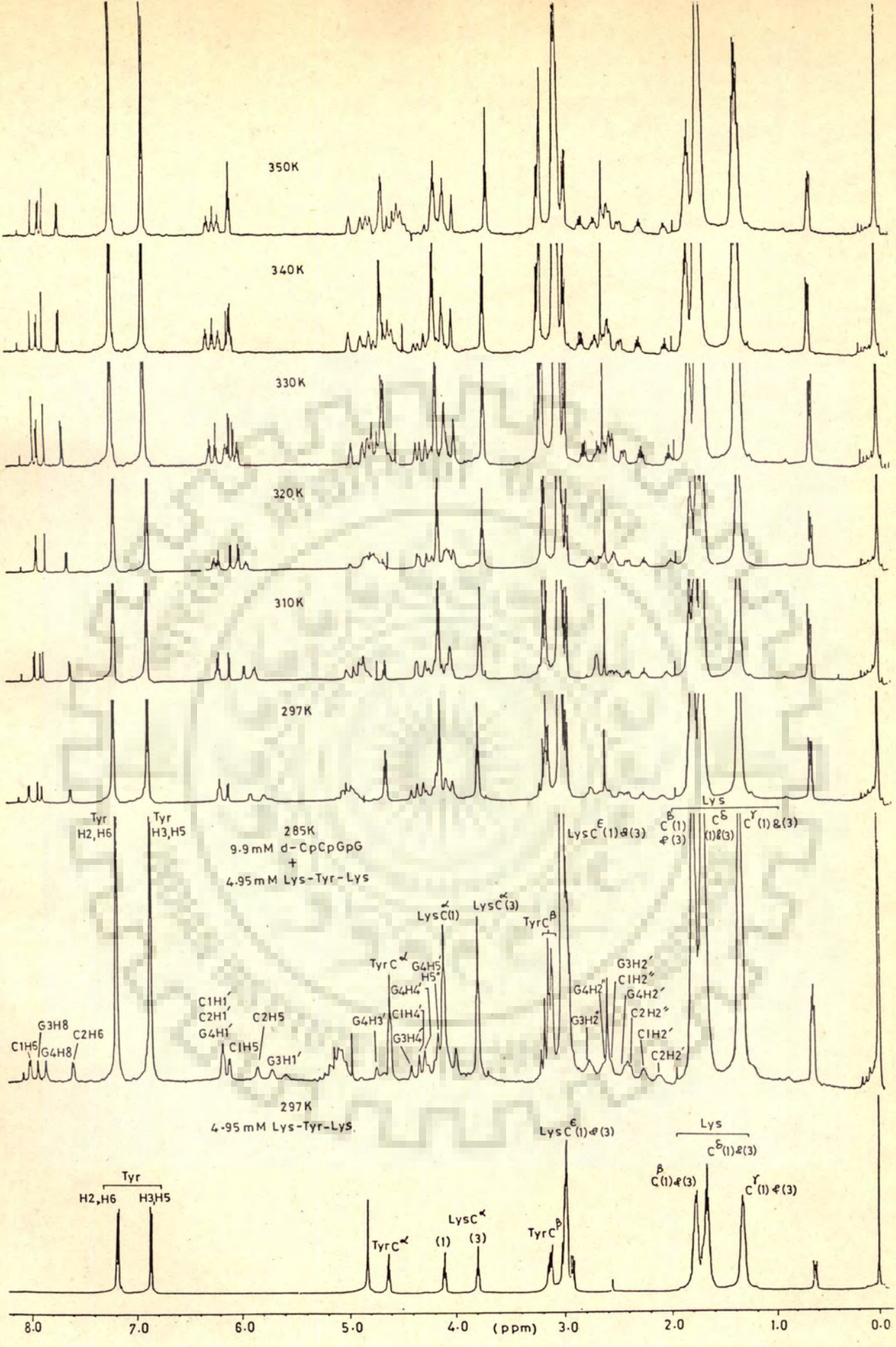


Fig. 6.7 500 MHz proton NMR spectra of a mixture of 9.9 mM d-CpCpGpG and 4.95 mM Lys-Tyr-Lys at different temperatures in D_2O (pH=7.0) containing 0.25 mM EDTA. For comparison 500 MHz spectrum of 4.95 mM Lys-Tyr-Lys in D_2O solution at 297 K is also shown. Ref.DSS.

Table 7.2 : Partitioning of stacking energies (Kcal/mole) for stacking of Trp, Tyr and Phe with bases and base pairs.

	Minimum E_{TOT} (Kumar & Govil) (78)	Minimum E_{TOT} (our results)	E_{el}	E_{pcl}	E_{disp}	E_{rep}	Distance for minimum energy configration (in Å)
G-Trp	-11.5	-7.99	-0.50	-1.10	-10.36	3.98	3.4
G-Tyr	-9.4	-9.98	-3.53	-1.37	-8.94	3.87	3.4
G-Phe	-9.1	-7.95	-1.96	-1.21	-9.07	4.29	3.4
C-Trp	-10.70	-9.46	-2.29	-1.95	-8.66	3.44	3.4
C-Tyr	-8.70	-8.83	-3.83	-0.93	-6.70	2.63	3.4
C-Phe	-8.30	-5.38	-0.11	-1.63	-5.28	1.64	3.4
GC-Trp	-14.5	-10.59	-1.14	-1.16	-12.23	3.95	3.4
GC-Tyr	-12.6	-8.94	-0.83	-1.04	-10.57	3.98	3.4
GC-Phe	-10.8	-7.20	-0.04	-0.58	-9.09	2.51	3.4
AT-Trp	-14.2	-8.63	-0.04	-0.33	-12.44	4.10	3.4
AT-Tyr	-12.1	-9.10	-2.31	-0.44	-9.33	2.99	3.4
AT-Phe	-10.0	-7.44	-0.84	-0.37	-9.62	3.02	3.4

Table 7.3 : Partitioning of stacking energies (kcal/mole) for intercalation of Trp, Tyr, Phe within base pairs C-G and G-C for different distances for $\alpha = 36°$

	Distance between C-G & G-C base pairs of model system d-CpG (in Å)	E_{TOT}	E_{el}	E_{pol}	E_{rep}	E_{disp}	Distance (d) for minimum configuration (in Å)
	6.2	-19.8044	-2.1994	-4.0973	-35.0888	21.5811	3.1
	6.4	-20.7680	-2.0051	-3.6432	-30.2398	15.1201	3.2
Trp.	6.6	-20.7268	-1.6009	-2.8907	-27.5455	11.3102	3.3
	6.8	-20.1024	-1.6122	-2.3566	-24.5483	8.4148	3.4
	7.0	-19.1261	-1.5178	-1.9885	-21.7932	6.1734	3.5
	6.2	-18.1434	-4.7330	-1.3672	-29.8901	17.8469	3.1
	6.4	-18.7698	-4.3451	-1.1243	-26.0213	12.7209	3.2
Tyr	6.6	-18.5588	-3.9909	-0.9393	-22.6977	9.0690	3.3
	6.8	-17.8685	-3.6663	-0.7976	-19.8676	6.4631	3.4
	7.0	-16.8823	-3.3627	-0.6637	-17.4470	4.5910	3.5
	6.2	-13.5524	-0.4029	-1.6493	-27.0019	15.5016	3.1
	6.4	-14.9472	-0.3936	-2.0198	-24.3432	11.8094	3.2
Phe	6.6	-14.7868	-1.1002	-1.1021	-21.9661	9.3817	3.3
	6.8	-14.4742	-1.0522	-0.9172	-19.3173	6.8125	3.4
	7.0	-13.9194	-0.9632	-0.7269	-17.2198	4.9910	3.5

Table 7.4 : Partitioning of stacking energies for Trp, Tyr and Phe intercalated between base pair C-G and G-C for relative rotations of bases, 36° and 61°

System	Angle of rotation	α = 36°				α = 61°				Distance (d) for minimum configuration (Å)
	E_{el}	E_{pol}	E_{disp}	E_{rep}	E_{TOT}	E_{el}	E_{pol}	E_{disp}	E_{rep}	E_{TOT}	
d-CpG + Trp	-2.0051	-3.6432	-30.2398	15.1201	-20.7680	-2.5473	-1.6507	-33.0261	17.0676	-20.1564	3.2
d-CpG + Tyr	-4.3451	-1.1243	-26.0213	12.7209	-18.7698	-3.6914	-0.8964	-25.3500	11.7233	-18.2146	3.2
d-CpG + Phe	-0.3936	-2.0198	-24.3432	11.8094	-14.9472	-1.3829	-0.7072	-24.9526	12.3214	-14.7214	3.2

decreases in the order Trp > Tyr > Phe for cytosine base but decreases in order Tyr > Trp > Phe for Guanine base. For both bases G and C, the electrostatic term is more attractive for Tyr than for Trp and Phe.

Figs. 7.3 and 7.4 show the overlap geometries of bases C and G with Trp, Tyr and Phe in the minimum energy conformation on stacking. In all molecules studied, it is seen that the overlap between planar ring of aromatic amino acid and base is only partial. The overlap decreases in the order Trp > Tyr > Phe for cytosine but decreases in order Tyr > Trp > Phe for guanine. The extent of overlap is in accordance with the stacking energies observed. The energy values and overlap geometries are comparable, in general, to those obtained by Kumar and Govil [78].

INTERACTION WITH BASE PAIRS

The stacking energies obtained with base pairs show features similar to that observed for bases. Energies are minimum for vertical separation of 3.4 Å. The dispersion energies are most dominant contributions to the total energy. For G-C base pair energy decreases in order Trp > Tyr > Phe whereas for A-T base pair energy decreases in the order Tyr > Trp > Phe. Out of two base pairs G-C yields greater stable conformation for Trp than that for A-T base pair. However, for Tyr and Phe, A-T yields more stable

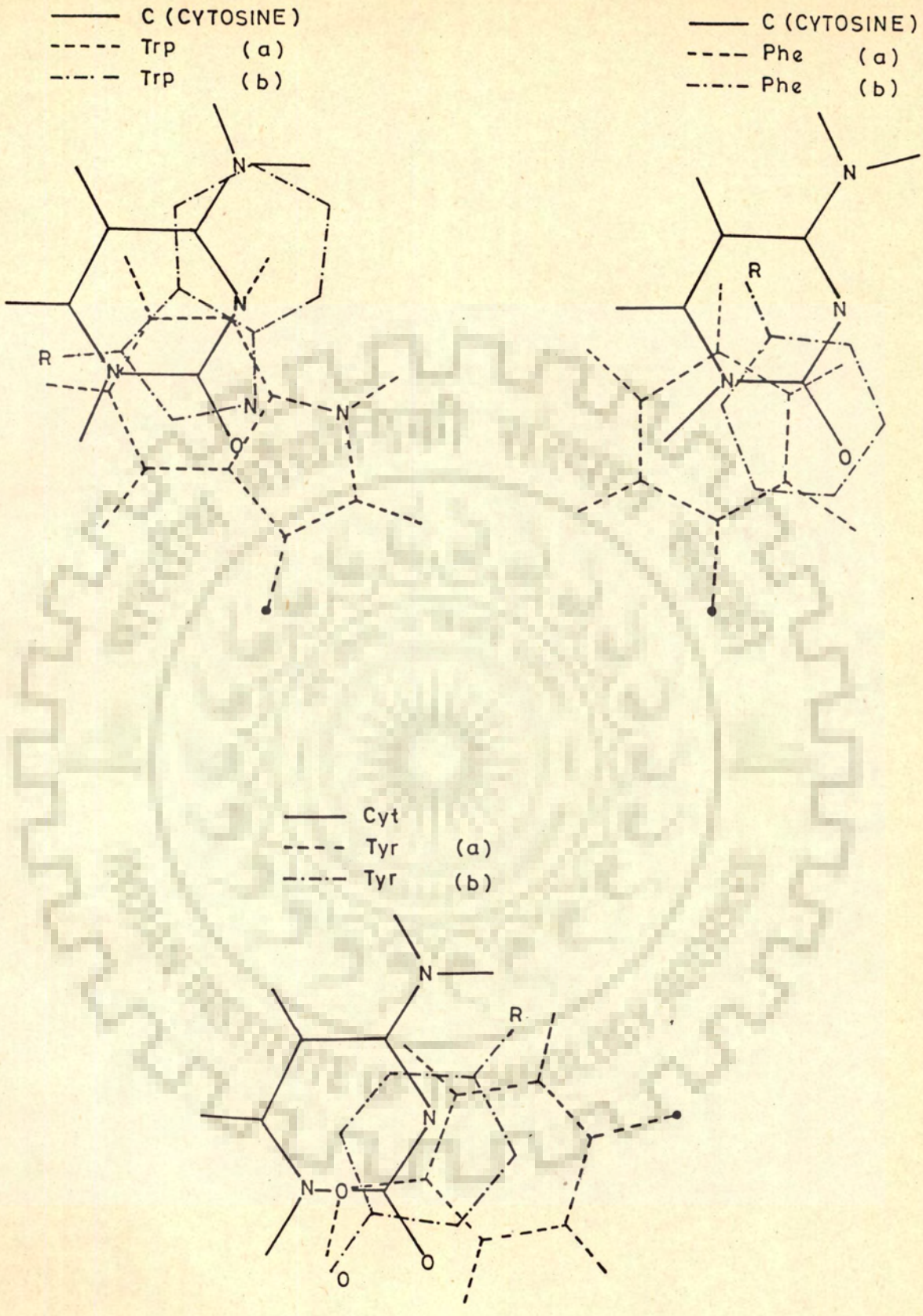


Fig. 7.3 Stacking of base C with Trp, Tyr and Phe **(a)** Our Results **(b)** Govil and coworkers

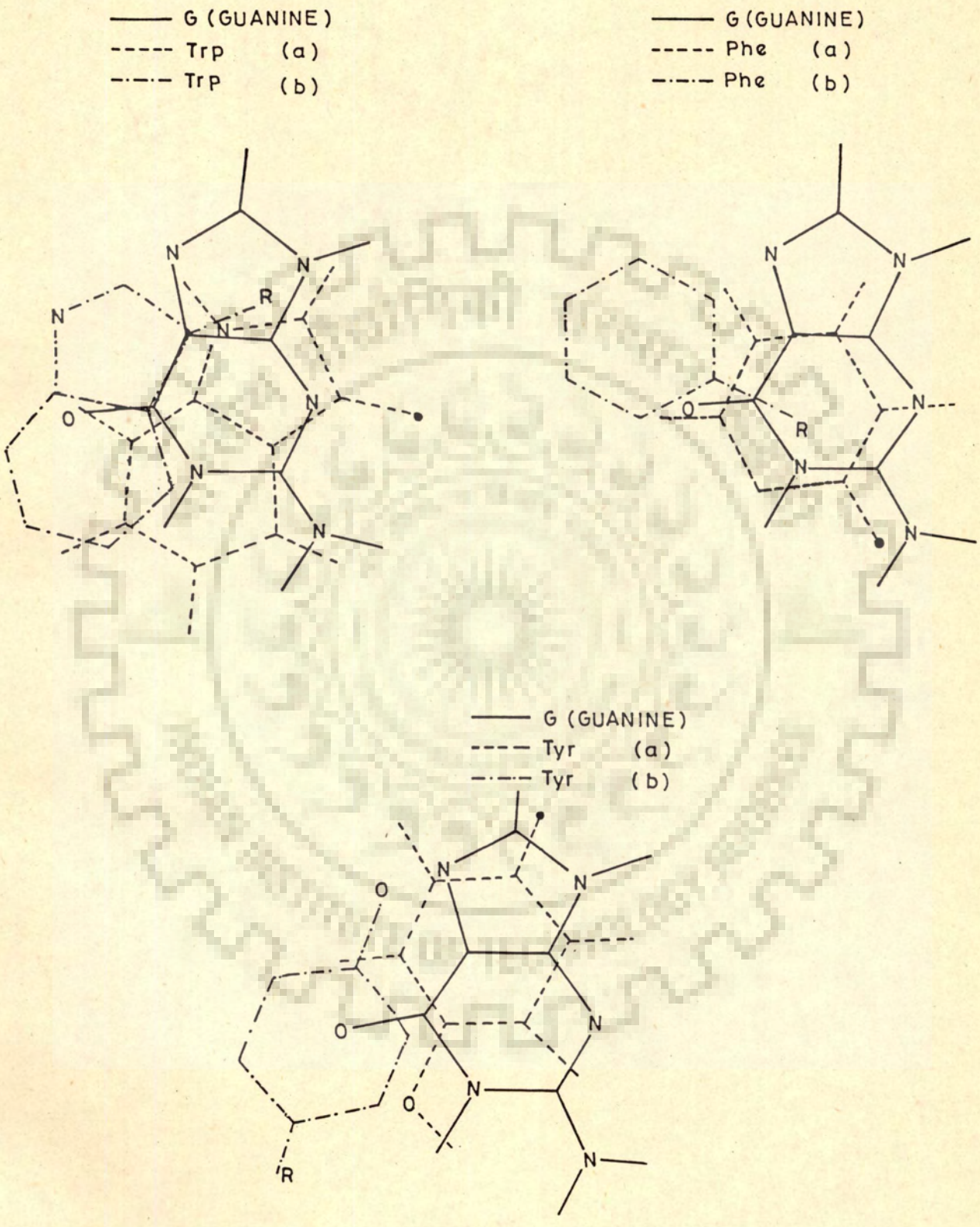


Fig. 7.4 Stacking of base G with Trp, Tyr and Phe **(a)** Our results **(b)** Govil and coworkers

complex than G-C base pair. The stacking geometries with base pairs (Figs. 7.5 and 7.6) show that overlap of Trp, Tyr and Phe occurs to a large extent with base pairs as compared to that with bases.

INTERACTION WITH C-G (ABOVE) AND G-C (BELOW) BASE PAIRS

A Dinucleotide d-CpG Helix Model System

Fig. 7.7 shows that the stacking energy is minimum when the distance between two base pairs is 6.4 Å. Figs. 7.8 and 7.9 show the overlap geometries for the optimised energy configurations obtained on intercalating Trp, Tyr and Phe between base pairs. It is seen that the overlap with base pair is substantial. The total stacking energy is almost double of that obtained for G-C base pair. The complex of Trp, Tyr and Phe are significantly more stable when they are intercalated between two base pairs, C-G and G-C, than that on stacking with a base pair G-C. The dispersion terms are the dominant ones.

The total energy apparently seems to be same for two sets of configurations (Table 7.4) in which the relative orientation angle between base pairs is 36° and 61°. However, the overlap geometries show that the intercalation occurs in a very different manner in two cases. For example, in case of Trp, the five membered ring of Trp, overlaps with cytosine of upper base pair C-G for α = 36° whereas part

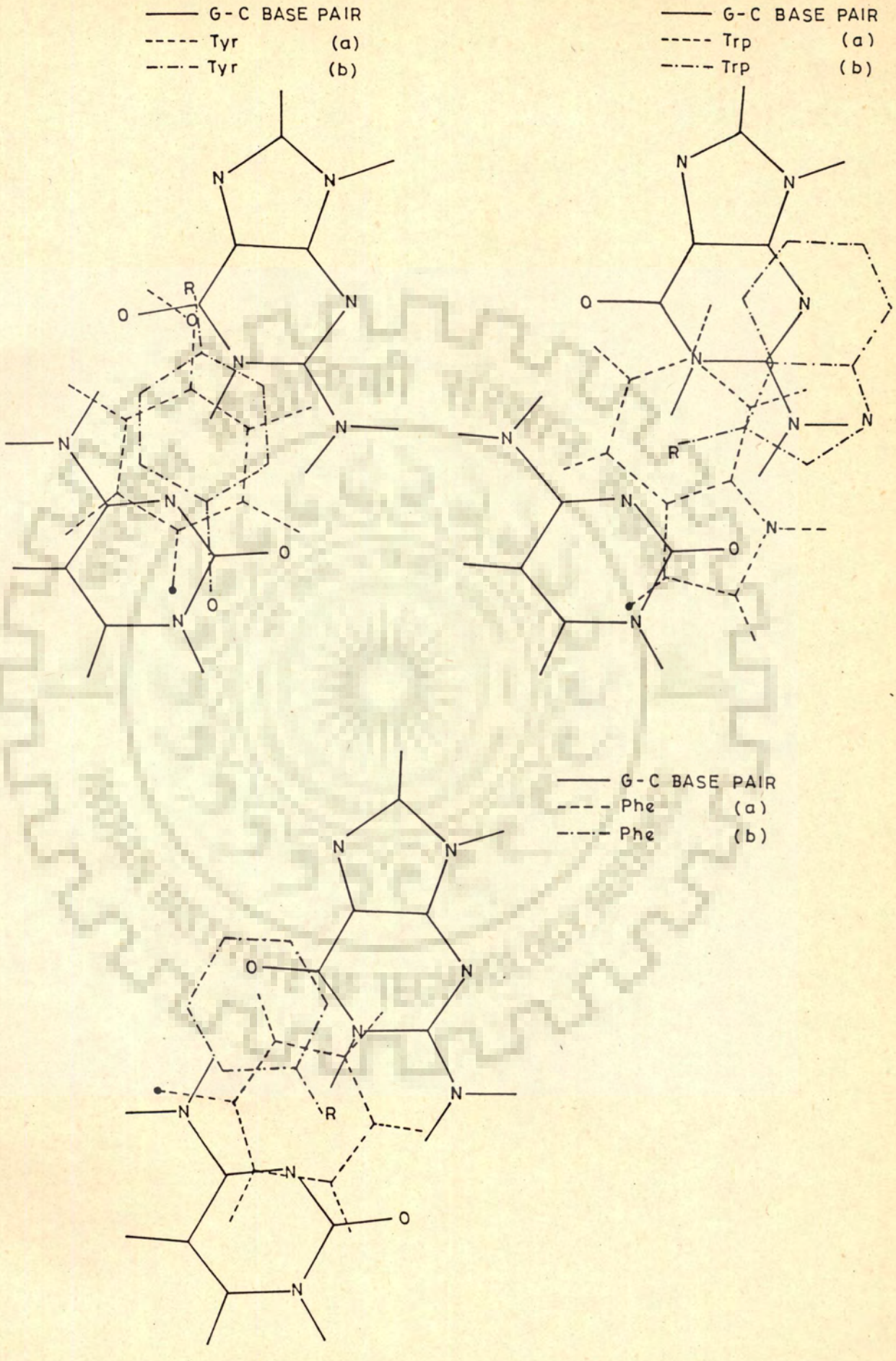


Fig. 7.5 Stacking of G-C base pair with Trp, Tyr and Phe **(a)** Our results **(b)** Govil and coworkers

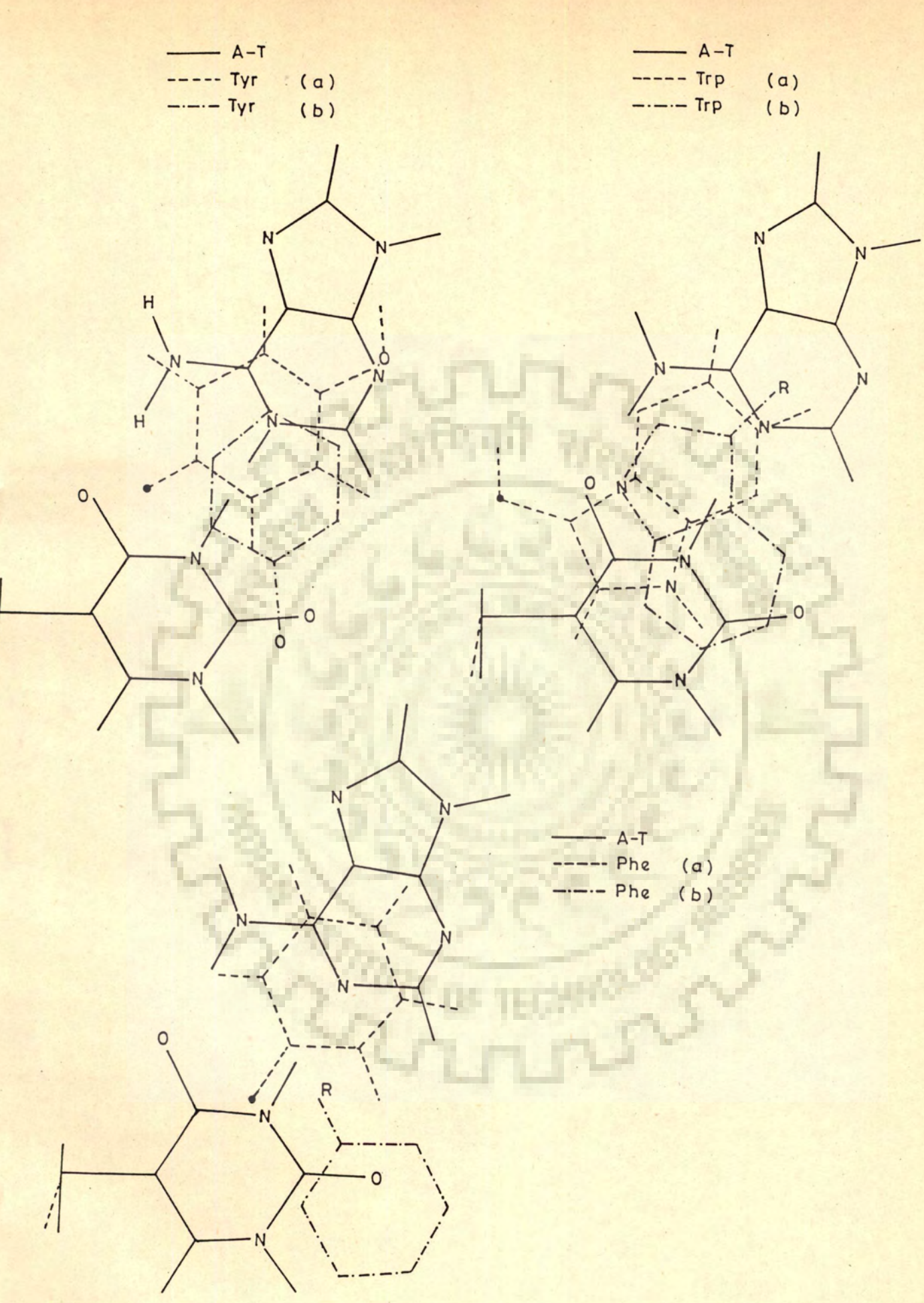


Fig. 7.6 Stacking of A-T base pair with Trp, Tyr and Phe **(a)** Our results **(b)** Govil and coworkers

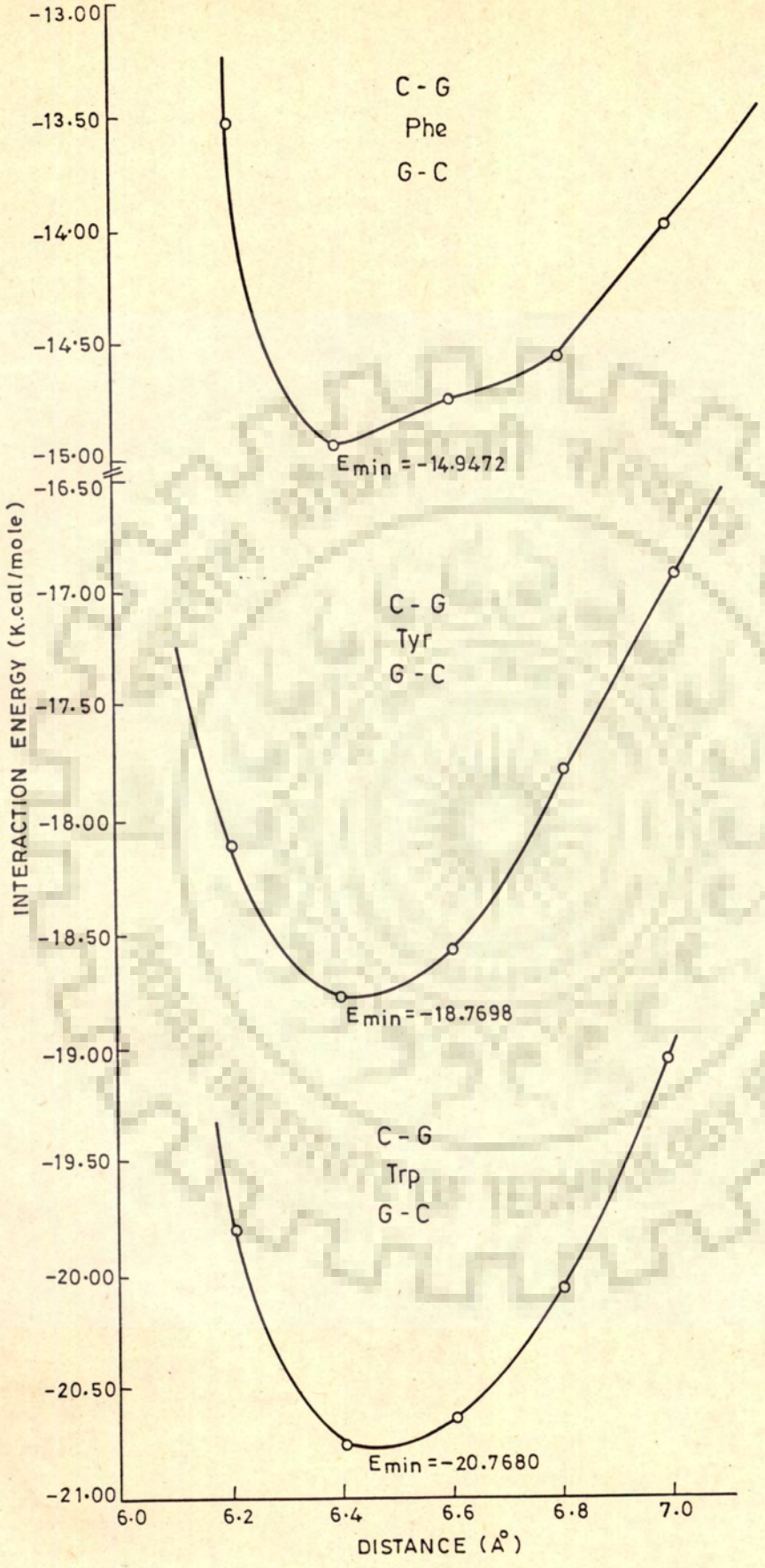


Fig. 7.7 Variation of stacking energy with distance of aromatic amino acids between two base pairs C-G and G-C for $\alpha = 36°$

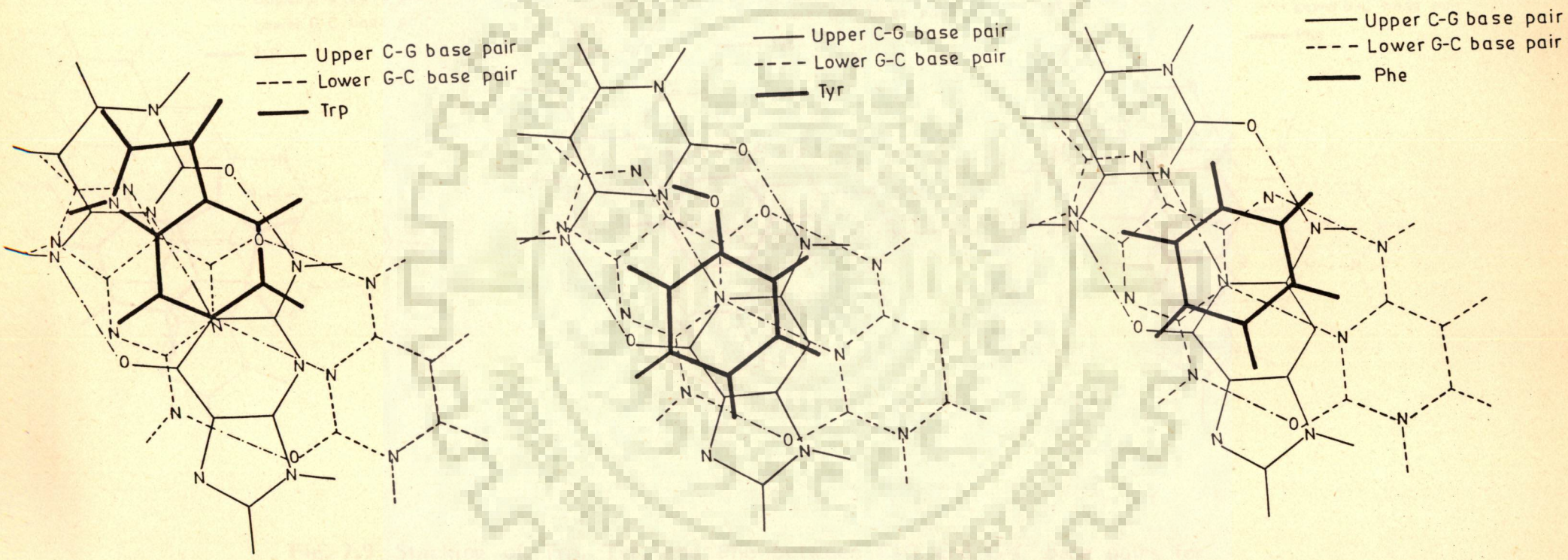


Fig. 7.8 Stacking of Trp, Tyr and Phe between C-G and G-C base pairs for $\alpha = 36°$

Table 8.3 : The results on binding of d-CpCpGpG to tripeptide containing Tyr

			d-CpCpGpG	d-CpCpGpG + Lys-Tyr-Lys
1.	Sugar Conformation	C1	01'-endo	01'-endo
		C2	01'-endo	01'-endo
		G3	01'-endo	01'-endo
		G4	01'-endo	01'-endo
2.	χ_{CN}	C1	anti	anti
		C2	anti	anti
		G3	anti	anti
		G4	anti	anti
3.	Helix sense		right-handed	right-handed
4.	$T_{\frac{1}{2}}$ (K)		310-312	310-320
5.	$\Delta\delta$(ppm) of ring protons of peptide on binding (up-field shifts)		285 K	.060-.065
			297 K	.044-.046
			320 K	.021-.022
6.	Binding to double-helix			more
7.	Binding to single-helix			less
8.	NOE's between ring protons of peptide and base protons of d-CpCpGpG			not seen

Table 8.3 : The results on binding of d-CpCpGpG to tripeptide containing Tyr

			d-CpCpGpG	d-CpCpGpG + Lys-Tyr-Lys
1.	Sugar Conformation	C1	01'-endo	01'-endo
		C2	01'-endo	01'-endo
		G3	01'-endo	01'-endo
		G4	01'-endo	01'-endo
2.	χ_{CN}	C1	anti	anti
		C2	anti	anti
		G3	anti	anti
		G4	anti	anti
3.	Helix sense		right-handed	right-handed
4.	$T_{\frac{1}{2}}$ (K)		310-312	310-320
5.	$\Delta\delta$(ppm) of ring protons of peptide on binding (up-field shifts)		285 K	.060-.065
			297 K	.044-.046
			320 K	.021-.022
6.	Binding to double-helix			more
7.	Binding to single-helix			less
8.	NOE's between ring protons of peptide and base protons of d-CpCpGpG			not seen

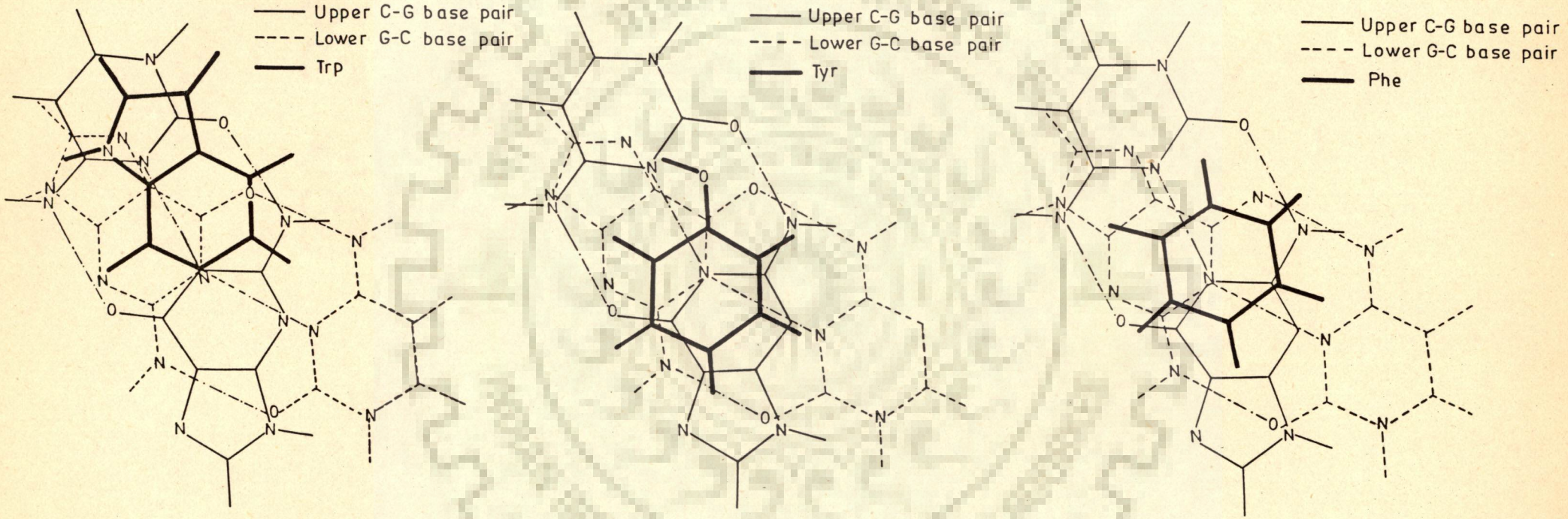


Fig. 7.8 Stacking of Trp, Tyr and Phe between C-G and G-C base pairs for $\alpha = 36°$

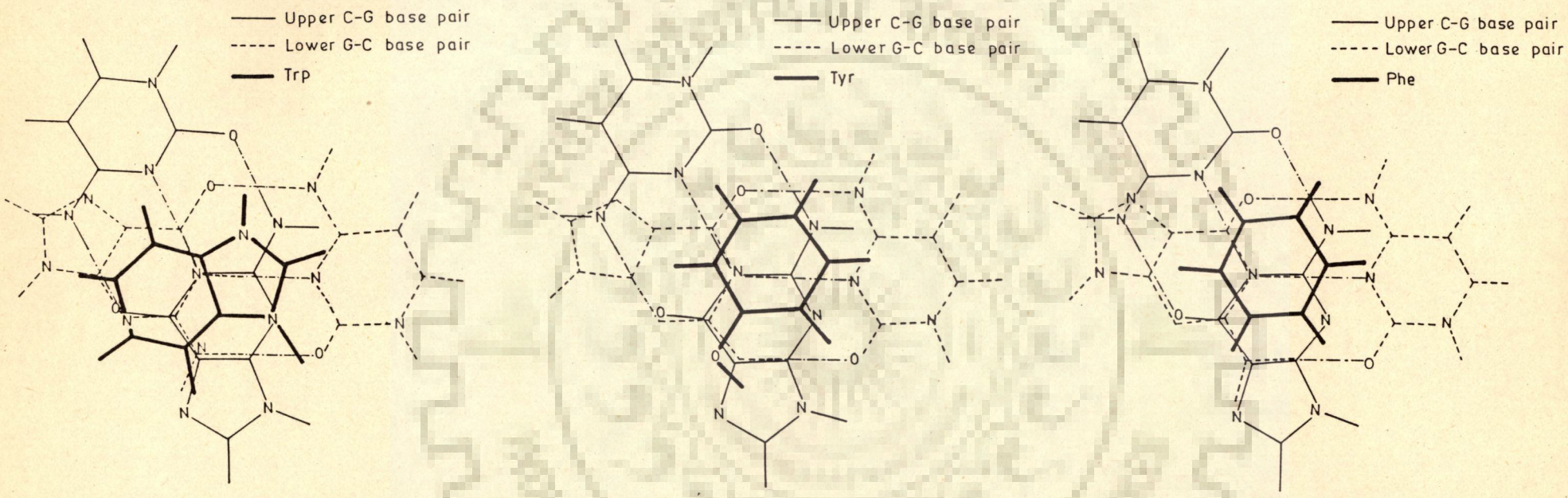


Fig. 7.9 Stacking of Trp, Tyr, and Phe between C-G and G-C base pairs for $\alpha = 61°$

of both five and six membered rings of Trp overlap with guanine of upper base pair C-G for $\alpha = 61°$. In both cases, interaction decreases in the order Trp > Tyr > Phe. Also the dispersion energy decreases in the same order.

SPECIFIC RECOGNITION BY STACKING INTERACTION

The association of aromatic amino acids with nucleic acids and their fragments have been studied by many experimental techniques and a specificity in recognition due to the stacking association has been suggested, directly or indirectly. It has been shown that solubility enhancement of guanosine [5] in presence of 0.05 M aromatic amino acids are 1.11, 0.33 and 0.13 for Trp, Tyr and Phe respectively. Binding constants of 5'-GMP with 1:1 copolymers of Lys and aromatic amino acids have been shown to be 14.7, 7.3 and 5.0 for Trp, Tyr and Phe, respectively [6]. Circular dichroism studies with Poly A have shown [87] that the interaction of aromatic amino acids decrease in the order Trp > Tyr > Phe. It has also been shown [117] that amides of Phe, Tyr, Trp at high concentration increase the Tm of double-stranded polyribonucleotides and the affinity for single-stranded polyribonucleotides decreases in the order Trp > Tyr > Phe. Some of the NMR studies by Helene and coworkers [33] and those discussed in Chapter-3 and 5 in the present thesis, also indicate preferential stacking. These results qualitatively support the theoretical predictions.

It may be noted that in addition to the geometries obtained from minimum energy values and shown in Figures, there are other configurations which have stacking energies within 0.5 kcal/mole of the optimised structure. This clearly suggests that a variety of active site geometries, either inherent or induced in the native systems can facilitate the binding of proteins to nucleic acids and a considerable conformational freedom exists in the stacked complexes. For this very reason it is not appropriate to give too much emphasis on exact total energy values and sketch it too far to predict the behaviour in native systems. However, one can not under estimate the trends in theoretical results obtained. They certainly act as guidelines and the diverse values of total stacking energies, electrostatic and dispersion terms and the overlap geometries have shown that there are specific trends in protein-nucleic acid interactions. Although stacking energies are characterised by dominant contributions from dispersion terms, the relative trends in stacking energies can not be explained by the variations in these contributions alone for bases and base pairs. However, for intercalation between two base pairs, wherein dispersion terms become much more significant, the variation in total stacking energy conforms with corresponding variation in dispersion terms e.g. Both E_{tot} and E_{disp} decreases in order Trp > Tyr > Phe.

CHAPTER–8

CONCLUSION

The results of binding of d-CpG to oligopeptides, Lys-Tyr-Lys, Lys-Trp-Gly-Lys and Lys-Phe-Lys are summarized in Table 8.1. It is clearly seen from the Table that the helix sense changes from double - helix at low temperature to single - helix at high temperatures. The upfield shifts in binding of d-CpG to various oligopeptides containing aromatic amino acids show that the aromatic rings of peptide stack or intercalate between the base pairs of d-CpG preferentially with double - helix. The NOE's of aromatic ring protons of the amino acid residues with base protons is a more direct proof of the same. Further the binding is maximum with Trp and decreases in the order of Trp > Tyr > Phe. The helix sense and glycosidic bond rotation do not change on binding to any of these peptides. The sugar conformation of cytosine base changes from O1'-endo to C2'-endo conformation on binding to Trp containing peptide, whereas on binding to other peptides containing Tyr and Phe, sugar conformations remain unchanged. Apparently it may be possible that stacking or intercalation of aromatic ring of peptide takes place between base pairs of d-CpG.

The results of binding of d-GpC to oligopeptides Lys-Tyr-Lys and Lys-Trp-Gly-Lys are summarized in Table 8.2. The upfield shift in the aromatic ring proton resonances on binding

Table 8.1 : The results on binding of d-CpG to Trp, Tyr and Phe containing oligopeptides

		d-CpG	d-CpG + Lys-Trp-Gly-Lys	d-CpG + Lys-Tyr-Lys	d-CpG + Lys-Phe-Lys
1. Sugar Conformation.	G	01'-endo	01'-endo	01'-endo	01'-endo
	C	01'-endo	C2'-endo	01'-endo	01'-endo
2. χ_{CN}	G	anti	anti	anti	-
	C	anti	anti	anti	-
3. Helix sense.		right-handed	right-handed	right-handed	
4. $T_{\frac{1}{2}}$(K)		295-297	295-297	305-322	
5. $\Delta\delta$(ppm) of ring protons of peptide on binding (upfield shifts).	277 K		.042-.054		
	285 K		0.028-.050	0.012-.017	
	297 K		0.014-0.030	0.010	.008
	335 K		-	0.004	-
6. Binding to double helix.		-	more	more	-
7. Binding to single helix.		-	less	less	-
8. NOE's between ring protons of peptide and base protons of d-CpG.			seen	seen	seen

Table 8.2 : The results on binding of d-GpC to Trp and Tyr containing oligopeptides

			d-GpC	d-GpC + Lys-Trp-Gly-Lys	d-GpC + Lys-Tyr-Lys
1.	Sugar Conformation	G	01'-endo	01'-endo	01'-endo
		C	01'-endo	01'-endo	01'-endo
2.	χ_{CN}	G		anti	anti
		C		anti	anti
3.	Helix sense			right-handed	right-handed
4.	$T_{\frac{1}{2}}$(K)		286	303-305	302-307
5.	$\Delta\delta$(ppm) of ring protons of peptide on binding (upfield shifts)		285 K	.029-.050	.011-.032
			297 K	.017-.038	.007-.022
			320 K	.004-.027	.000-.010
6.	Binding to double-helix			more	more
7.	Binding to single-helix			less	less
8.	NOE's between ring protons of peptide and base protons of d-GpC			seen	seen

of d-GpC to peptides clearly indicate that the aromatic ring of peptide is in close proximity to the dinucleotide. This change decreases with increase in temperature and binding seems to be preferential to double - helix. NOE's of ring protons with base protons give evidence in support of this. Also the binding is more with Trp than with Tyr. The sugar conformation, helix sense and glycosidic bond rotation do not show any change on binding to any of these peptides.

In case of d-GpC binding to Lys-Trp-Gly-Lys, the ring proton resonances of the Trp shift upfield by about .017-.038 ppm at 297 K which is more than the corresponding change observed on binding of Lys-Trp-Gly-Lys to d-CpG (.014-.030). This clearly shows that the binding of Lys-Trp-Gly-Lys with d-GpC is stronger than with d-CpG.

Table 8.3 summarizes the results of binding of Lys-Tyr-Lys to d-CpCpGpG. The upfield shift in Tyr ring proton resonances is less and decreases with increase in temperature. This suggests that the binding is due to stacking/partial stacking/intercalation of Tyr ring between base pairs of d-CpCpGpG preferentially with double-stranded tetranucleotide. The helix sense, glycosidic bond rotation and sugar conformation do not change on binding to Lys- Tyr-Lys.

The results of binding of d-GpCpGpC with Lys-Tyr-Lys are tabulated in Table 8.4.The large upfield shift in Tyr ring proton resonances is obviously due to stacking of Tyr ring

Table 8.4 : The results on binding of d-GpCpGpC to tripeptide containing Tyr

			d-GpCpGpC	d-GpCpGpC+Lys-Tyr-Lys
1.	Sugar Conformation	G1	O1'-endo	O1'-endo
		C2	O1'-endo	O1'-endo
		G3	O1'-endo	O1'-endo
		C4	O1'-endo	O1'-endo
2.	χ_{CN}	C1	anti	anti
		C2	high anti	anti/high anti
		G3	anti	anti
		C4	high anti	anti/high anti
3.	Helix sense		right-handed	right-handed
4.	$T_{\frac{1}{2}}$ (K)		314	319-324
5.	$\Delta\delta$(ppm) of ring protons of peptide on binding(up-field shifts)	285 K	—	.140-.141
		297 K	—	.108-.106
		315 K	—	.066-.061
6.	Binding to double-helix			more
7.	Binding to single-helix			less
8.	NOE's between ring protons of peptide and base protons of d-GpCpGpC			seen

between base pairs of d-GpCpGpC. Further, the change in chemical shift decreases with increase in temperature indicating that the binding preferentially occurs to double-stranded d-GpCpGpC. On complex formation, sugar conformation, glycosidic bond rotation and helix sense remain unchanged. However, the appearance of intermolecular NOE's between Tyr ring protons and sugar protons indicate proximity of tripeptide to tetranucleotide in the complex. On comparing binding of Lys-Tyr-Lys to various oligonucleotides d-CpG, d-GpC, d-CpCpGpG and d-GpCpGpC, it becomes clear that the upfield chemical shift observed for Tyr ring proton resonances is less ($\Delta\delta \sim .06$ ppm, 285 K) in d-CpCpGpG than that to d-GpCpGpC ($\Delta\delta \sim .14$ ppm, 285 K). Also it may be clearly seen that the changes in chemical shift of Tyr ring proton resonances on binding to d-CpCpGpG ($\Delta\delta \sim .06$ ppm, 285 K) is to a larger extent than that to d-CpG ($\Delta\delta \sim .015$ ppm, 285 K). Hence it can be concluded that the binding of Lys-Tyr-Lys is maximum with d-GpCpGpC and decreases in the order d-GpCpGpC > d-CpCpGpG > d-GpC > d-CpG. The binding with d-GpCpGpC is more with tripeptide which is obviously due to the fact that this tetranucleotide contains two GpC sites and one CpG site as compared to d-CpCpGpG which has only one CpG site. Thus it can be inferred that GpC site is more preferable for binding with peptides than CpG site.

The results of the theoretical studies (Table 7.1) results clearly show that for intercalation of aromatic amino acids in between base pairs C-G and G-C, the binding is maximum with Trp and decreases in the order Trp > Tyr > Phe which is in agreement with the NMR results that we have obtained.

Hence these results clearly establish the preference for GpC site over CpG site which indicates that there is specificity for base sequence. Such binding to preferential base sequence is likely to have important implications in specific recognition of nucleic acid by proteins.

REFERENCES

1. Alexander, M.E., Burgum, A.A., Noall, R.A., Shaw, M.D. and Matthews, K.S., (1977) Modification of tyrosine residues of the lactose repressor protein. Biochim. Biophys. Acta. 493, 367-379.

2. Alma, N.C.M., Harmsen, B.J.M., Hilbers , C.W., Van der Marel, C.W. and Van Boom, J.H. (1981) 500 MHz ^{1}H NMR study of the role of lysines and arginines in the binding of gene 5 protein to oligoadenylic acids. FEBS Lett. 135, 15-20.; Alma, N.C.M., Harmsen, B.J.M., Hull, W.E., Van der Marel, G., Van Boom , J.H. and Hilbers, C.W. (1981) Double resonance experiments at 500 MHz on gene 5 protein and its complex with octadeoxyriboadenylic acid. Biochemistry. 20, 4419-4428.

3. Altona, C. and Sundaralingam, M. (1973) Conformational analysis of the sugar ring in nucleosides and nucleotides improved method of interpretation of proton magnetic resonance coupling constants. J. Amer. Chem. Soc. 95, 2333-2344.

4. Anderson, W.F., Ohlendorf, D.H., Takeda, Y. and Matthews, B.W. (1981) Structure of the cro repressor from bacteriophage λ and its interaction with DNA. Nature. 290, 754-758.

5. Arcaya, G., Pantoja, M.E., Pieber, M., Romero, C. and Toha, J.C. (1971) Z. Naturforsch, Tiel. B 26, 1026-1030.

Cited in Kumar, N.V. and Govil, G. (1984) Theoretical Studies on protein-nucleic acid interactions. III. Stacking of aromatic amino acids with bases and base pairs of nucleic acids. Biopolymers. 23, 2009-2024.

6. Arfmann, H.A., Labitzke, R., Lawaczeck, R. and Wagner, K.G. (1974) Aromatic amino acid-lysine Copolymers. Conformation and specificity of nucleotide interaction. Biochimie. 56, 53-60.

7. Arnott, S. and Hukins, D.W.L. (1973) Refinement of the structure of B-DNA and implications for the analysis of x-ray diffraction data from fibres of biopolymers. J. Mol. Biol. 81, 93-105.

8. Aue, W.P., Bartholdi, E. and Ernst, R.R. (1976) Two-dimensional spectroscopy. Application to nuclear magnetic resonance. J. Chem. Phys. 64, 2229-2246.

9. Bax, A. and Freeman, R. (1981) Investigation of complex networks of spin-spin coupling by two-dimensional NMR. J. Magn. Reson. 44, 542-561.

10. Bell, R.A. and Saunders, J.K. (1973) In topics in stereochemistry. (Allinger, N.L. and Eliel, E.L. eds.) 7, pp. 1-92, John Wiley. New York.

11. Berg, O.G., Winter, R.B. and Von Hippel, P.H. (1982) How do genome regulatory proteins locate their DNA target sites. Trends Biochem. Sci. 7, 52-55.

12. Bloch, F., Hansen, W.W. and Packard,M. (1946) Nuclear induction. Phys. Rev. 69, 127. and Purcell, E.M. Torrey, H.C., Pound, R.V. (1946) Resonance absorption by nuclear magnetic moments in a solid. Phys. Rev. 69, 37-38.

13. Bloomer, A.C., Champness, J.N., Bricongne, G., Staden, R. and Klug, A. (1978) Protein disk of tobacco mosaic virus at 2.8Å resolution showing the interactions within and between subunits. Nature. 276, 362-368.

14. Bram, S., Butler-Browns, G., Bradbury, E.M., Baldwin, J.P., Reiss, C. and Ibel, K. (1974) Biochimie. 56, 987. Cited in 'Kumar, N.V. (1983) Ph.D. Thesis'. Tata Institute of Fundamental Research, Bombay, India.; Chinsky, L. and Turpin, P.Y. (1982) Poly (dA-dT) Complexes with Histone H1 and pancreatic ribonuclease specific base recognition evidenced by ultraviolet resonance Raman spectroscopy. Biopolymers. 21, 277-286.

15. Broido, M.S., James, T.L., Zon, G. and Keepers, J.W. (1985) Investigation of the solution structure of a DNA octamer $[d(GGAATTCC)]_2$ using two-dimensional nuclear Overhauser enhancement spectroscopy Eur. J. Biochem. 150, 117-128.

16. Bruskov, V.I. (1975), The recognition of nucleic acid bases by amino acids and peptides with the aid of hydrogen bonds. Molekulyarnaya Biologiya. 9, 304-309, Seeman, N.C., Rosenberg, J.M. and Rich, A. (1976) Sequence specific recognition of double-helical nucleic acids by pro-

teins. Proc. Natl Acad. Sci., U.S.A. 73, 804-808.

17. Buck, F., Ruterjans., H., Kaptein, R. and Beyreuther, K. (1980) Photochemically induced dynamic nuclear polarization investigation of complex formation of the NH_2-terminal DNA binding domain of lac repressor with poly d(AT). Proc.Natl Acad. Sci. U.S.A. 77, 5145-5148.

18. Caillet, J. and Claverie, P. (1975) Theoretical evaluation of the intermolecular interaction energy of a crystal: Application to the analysis of crystal geometry. Acta Cryst. A31, 448-461.

19. Cantor, C.R. and Schimmel, P.R. (1980) In 'Biophysical Chemistry. Part II : Techniques for the study of biological structure and function'. W.H. Freeman, New York.

20. Caruthers, M.H. (1980) Deciphering the protein-DNA recognition code. Acc . Chem. Res. 13, 155-160.

21. Casas-Finet, J.R., Toulme, J.J., Cazenave, C. and Santus, R. (1984) Role of tryptophan and cysteine in the binding of gene 32 protein from phage T_4 to single-stranded DNA. Modification of crucial residues by oxidation with selective free-radical anions. Biochemistry. 23, 1208-1213.

22. Charlier, M., Maurizot, J.C. and Zaccai, G. (1980) Neutron scattering studies of the lac repressor. Nature. 286, 423-425.

23. Chary, K.V.R., Hosur, R.V. and Govil, G. (1987) Novel solution conformation of DNA observed in d(GAATTCGAATTC) by two-dimensional NMR spectroscopy. Biochemistry. 26, 1315-1322.

24. Chen, K.I., Gresh, N. and Pullman, B. (1987) A theoretical exploration of conformational aspects of Ethidium Bromide intercalation into a $d(CpG)_2$ minihelix. Biopolymers. 26, 831-848.

25. Claverie, P. (1978) Elaboration of approximate formulas for the interactions between large molecules. Application in organic chemistry. In 'Intermolecular interactions : from Diatomics to Biopolymers'. (Pullman, B., ed.) pp. 69-305, John Wiley, New York.

26. Coleman, J.E., Anderson, R.A., Ratcliffe, R.G. and Armitage, I.M. (1976) Structure of gene 5 protein oligonucleotide complexes as determined by ^{1}H, ^{19}F, and ^{21}P nuclear magnetic resonance. Biochemistry. 15, 5419-5430.

27. Coleman, J.E., and Armitage, I.M. (1978) Tyrosyl-base-phenylalanyl intercalation in gene 5 protein-DNA complexes: proton nuclear magnetic resonance of selectively deuterated gene-5 protein. Biochemistry. 17, 5038-5045.

28. Collins, D.M., Cotton, F.A., Hazen, E.E., Jr. and Legg, M.J. (1975) I: The nucleotide binding site of Staphylococcal nuclease and II : A new approach to the refinement of

the crystal structures of biological macromolecules. In 'structures and conformation of Nucleic acids and Protein -Nucleic acid Interactions' (Sundaralingam, M. and Rao, S.T., eds.) pp 317-331.

29. Cotton, F.A., Hazen, E.E., Jr., Day , V.W., Larsen, S., Norman, J.G., Jr. Wong, S.T.K. and Johnson, K.H. (1973) Biochemical importance of the binding of phosphate by arginyl groups. Model compounds containing methylguanidinium ion. J. Amer. Chem. Soc. 95, 2367-2369; Cotton, F.A., Day, V.W., Hazen, E.E., Jr. and Larsen, S. (1973) Structure of methylguanidinium dihydrogenorthophosphate. A model compound for arginine-phosphate hydrogen bonding. J. Amer. Chem. Soc. 95, 4834-4840.

30. Cotton, F.A., Hazen, E.E., Jr., and Legg, M.J. (1979) Staphylococcal nuclease : proposed mechanism of action based on structure of enzyme thymidine 3'-5'-biphosphate-calcium ion complex at 1.5 Å resolution. Proc. Natl Acad. Sci., U.S.A. 76, 2551-2555.

31. Cruse, W.B.T., Egert, E., Kennard, O., Sala, G.B., Salisburg, S.A. and Viswamitra, M.A. (1983) Self base pairing in a complementary deoxydinucleoside monophosphate duplex : crystal and molecular structure of deoxycytidylyl-(3'-5')-deoxyguanosine. Biochemistry. 22, 1833-1839.

32. Cuatrecasas, P., Edelhoch, H. and Anfinsen, C.B. (1967) Fluorescence studies of the interaction of nucleotides

with the active site of the nuclease of staphylococcus aureus. Proc. Natl Acad. Sci., U.S.A. 58, 2043-2050.

33. Dimicoli, J.L. and Helene, C. (1974) Interactions of aromatic residues of proteins with nucleic acids. I. Proton magnetic resonance studies of the binding of tryptamine and tryptophan containing peptides to poly (adenylic acid) and deoxyribonucleic acid. Biochemistry. 13, 714-723; Dimicoli, J.L. and Helene, C. (1974) Interactions of aromatic residues of proteins with nucleic acids. II. Proton magnetic resonance studies of the binding of tyramine and tyrosine containing peptides to poly (adenylic acid) and deoxyribonucleic acid. Biochemistry. 13, 724-730.

34. Doan, T.L., Toulme, J.J. and Helene, C. (1984) Involvement of tryptophyl residues in the binding of model peptides and gene 32 protein from phage T_4 to single-stranded polynucleotides. A spectroscopic method for detection of tryptophan in the vicinity of nucleic acid bases. Biochemistry. 23, 1202-1207.

35. Draper, D.E. and Von Hippel, P.H. (1978) Nucleic acid binding properties of Escherichia Coli ribosomal protein S1. II. cooperativity and specificity of binding site II. J. Mol. Biol. 122, 339-359.

36. Eddington, P. and Harding, M.M. (1974) The crystal structure of DL-Histidine. Acta Cryst. B30, 204-206.

37. Fanning, T.G. (1975) Iodination of Escherichia Coli lac repressor. Effect of tyrosine modification on repressor activity. Biochemistry. 14, 2512-2520.

38. Feigon, J. Wright, J.M., Leupin, W. Denny, W.A. and Kearns, D.R. (1982) Use of two-dimensional NMR in the study of a double stranded DNA decamer. J. Amer. Chem. Soc. 104, 5540-5541; Feigon, J., Denny, W.A., Leupin, W. and Kearns, D.R. (1983) Proton nuclear magnetic resonance investigation of the conformation and dynamics in the synthetic deoxyribonucleic acid decamers d(ATATCGATAT) and d(ATATGCATAT). Biochemistry. 22, 5930-5942; Feigon, J., Leupin, W., Denny, W.A. and Kearns, D.R. (1983) Two-dimensional proton nuclear magnetic resonance investigation of the synthetic deoxyribonucleic acid decamer $d(ATATCGATAT)_2$. Biochemistry. 22, 5943-5951.

39. Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M. and Klug, A. (1977) Structure of nucleosome core particles of chromatin. Nature. 269, 29-36; Crothers, D.M., Dattagupta, N., Hogan, M., Klevan, L. and Lee, K.S.C. (1978) Transient electric dichroism studies of nucleosomal particles. Biochemistry. 17, 4525-4533,; Feigon, J. and Kearns, D.R. (1979) ^1H NMR investigation of the conformational states of DNA in nucleosome core particles. Nucl. Acids Res. 6, 2327-2337.

40. Frechet, D. Cheng, D.M., Kan. L.S. and Tso, P.O.P. (1983)

Nuclear overhauser effect as a tool for the complete assignment of nonexchangeable proton resonances in short deoxyribonucleic acid helices. Biochemistry. 22, 5194-5200.

41. Frey. M.N., Koetzle, T.F., Lehmann, M.S. and Hamilton, W.C. (1973) Precision neutron diffraction structure determination of protein and nucleic acid components. X. A comparison between the crystal and molecular structure of L-tyrosine and L-tyrosine hydrochloride. J. Chem. Phys. 58, 2547-2556.

42. Gabbay, E.J., Sanford, K. and Baxter, C.S. (1972) Specific interaction of peptides with nucleic acids. Biochemistry. 11, 3429-3435.; Gabbay, E.J., Sanford, K., Baxter, C.S. and Kapicak, L. (1973) Specific interaction of peptides with nucleic acids. Evidence for a 'selective book mark' recognition hypothesis. Biochemistry. 12, 4021-4029.

43. Garssen, G.J. Hilbers, C.W., Schoenmakers, J.G.G. and Van Boom, J.H. (1977) Studies on DNA unwinding. Eur. J. Biochem. 81, 453-463.; 43b. Garssen, G.J., Kaptein, R., Schoemakers,J.G.G. and Hilbers, C.W. (1978) A photo CIDNP study of the interaction of oligonucoeotides with gene 5 protein of bacteriophage M 13. Proc. Natl Acad. Sci. U.S.A. 75, 5281-5285.

44. Geisler, N. and Weber, K. (1977) Isolation of the amino-terminal fragment of lactose repressor necessary for

DNA binding. Biochemistry. 16, 938-943.

45. George, J. Blakesley, R.W. and Chirikjian, J.G. (1980) Sequence-specific endonuclease Bam HI. J. Biol. Chem. 255, 6521-6524.

46. Giege, R., Moras, D. and Thierry, J.C. (1977) Yeast transfer RNA^{Asp} : A new high resolution X-ray diffracting crystal form of a transfer RNA. J. Mol. Biol. 115, 91-96,; Giege, R. Lorber, B., Ebel, J.P., Moras, D. and Thierry, J.C. (1980) Cristallisation due complexe forme' entre l'aspartate - t RNA de levure et son aminoacyl-tRNA synthetase specifique. C.R. Acad. Sci. Paris. 2, 393-396,; Moras, D. Comarmond, M.B., Fisher, J., Weiss, R., Thierry, J.C., Ebel, J.P. and Giege, R. (1980) Crystal structure of yeast $tRNA^{Asp}$. Nature. 288, 669-674.

47. Giessner-Prettre, C. and Pullman B. (1970) Intermolecular nuclear shielding values for protons of purines and flavins. J. theor. biol. 27, 87-95.; Giessner-Prettre, C. and Pullman, B. (1976) On the atomic or 'Local' contributions to proton chemical shifts due to the anisotropy of the diamagnetic susceptibility of the nucleic acid bases. Biochem. Biophys. Res. Com. 70, 578-581; Giessner-Prettre, C. and Pullman, B. (1981) On the atomic or 'Local' contributions to chemical shifts due to the anisotropy of the diamagnetic susceptibility of the aromatic side chain of amino acids and of the porphyrin ring. Biochem.

Biophys. Res. Com. 101, 921-926.

48. Goeddel, D.V., Yansura, D.G., and Caruthers, M.H. (1978) How lac repressor recognizes lac operator. Proc. Natl Acad. Sci. U.S.A. 75, 3578-3582.

49. Gopalakrishnan, B. and Bansal, M. (1985) Comparison of interproton distances in DNA models with nuclear overhauser enhancement data. J. Biosci. 8, 603-614.

50. Govil, G. (1976) Conformational structure of polynucleotides around the O-P bonds: Refined parameters for CPF calculations. Biopolymers. 15, 2303-2307.

51. Govil, G. and Hosur, R.V. (1982) In 'conformation of biological molecules' : New results from NMR' (Diehl, P., Fluck, E. and Kosfeld, R., eds.) 20, Springer-Verlag, New York.

52. Govil, G.,(Unpublished results)

53. Gresh, N. and Pullman, B. (1979) A theoretical study of the interaction of ammonium and guanidinium ions with the phosphodiester linkage. Theor. Chim. Acta. 52, 67-73.

54. Gronenborn, A.M. and Clore, G.M. (1985) Investigation of the solution structure of short nucleic acid fragments by means of nuclear Overhauser enhancement measurements.

Prog. In NMR spectroscopy 17, 1-32.

55. Gurskaya, G.V. (1965) Crystal structure of L-phenylalanine hydrochloride. Refinement and discussion of structure. Sov. Phys. Crystallogr. 9, 709-714.

56. Hare, D.R., Wemmer, D.E., Chou, S.H. and Drobny, G. (1983) Assignment of the non-exchangeable proton resonances of d(C-G-C-G-A-A-T-T-C-G-C-G) using two dimensional nuclear magnetic resonance methods. J. Mol. Biol. 171, 319-336.

57. Helene, C., Brun, F. and Yaniv, M. (1971) Fluorescence studies of interactions between Escherichia coli valyl-tRNA sysnthetase and its substrates., J. Mol. Biol. 58, 349-365; Helene, C., Montenay-Garestier, T. and Dimicoli, J.L. (1971) Interaction of tyrosine and tyramine with nucleic acids and their components. Fluorescence, nuclear magnetic resonance, and circular dichroism studies. Biochim. Biophys. Acta. 254, 349-365; Helene, C., Dimicoli, J.L. and Brun, F. (1971) Binding of tryptamine and 5-hydroxytryptamine (Serotonin) to nucleic acids. Fluorescence and proton magnetic resonance studies. Biochemistry 10, 3802-3809.

58. Helene, C. and Dimicoli, J.L. (1972) Interaction of oligopeptides containing aromatic amino acids with nucleic acids. Fluorescence and proton magnetic resonance studies. FEBS Lett. 26, 6-10.

59. Helene, C. (1975) Metal ion-mediated specific interactions between nucleic acid bases of polynucleotides and amino acid side chains of polypeptides. Nucl. Acids Res. 2, 961-969 ; Bere, A. and Helene, C. (1979) Formation of ternary complexes involving zinc or copper ions polynucleotides, and polypeptides containing glutamic acid and tyrosine residues. Biopolymers. 18, 2659-2672.

60. Helene, C. (1976) Specific recognition of nucleic acids by proteins. Studia Biophysica. 57, 211-222.

61. Helene, C. (1977) Specific recognition of guanine bases in protein-nucleic acid complexes. FEBS Lett. 74, 10-13.

62. Helene, C. and Lancelot, G. (1982) Interactions between functional groups in protein-nucleic acid associations. Prog. Biophys. Molec. Biol. 39, 1-68.

63. Holmes, K.C. (1980) Protein—RNA interactions during the assembly of tobacco mosaic virus. Trends Biochem. Sci. 5, 4-7.

64. Hosur, R.V. (1980) Role of protein backbone in specific recognition of nucleic acid base sequences : a hypothesis. Current science. 49, 928-931.

65. Hosur, R.V., Kumar,N.V. and Govil, G. (1981) Protein-nucleic acid interactions : investigations on the peptide

backbone interaction with polynucleotides. Int. J. Quant. Chem. 20, 23-32.

66. Hosur, R.V., Kumar, M.R., Roy, K.B., Zu-Kun, T., Miles, H.T. and Govil, G. (1985) In magnetic resonance in biology and medicine. (Govil, G.Khetrapal , C.L. and Saran, A., eds.) pp 243-260. Tata McGraw Hill, New Delhi.; Scheek, R.M., Boelens, R., Russo, N., Van Boom, J.H. and Kaptein, R. (1984) Sequential resonance assignments in ^{1}H NMR spectra of liogonucleotides by two-dimensional NMR spectroscopy. Biochemistry. 23, 1371-1376.

67. Jain, S.C. and Sobell, H.M. (1972) Stereochemistry of actinomycin binding to DNA I. Refinement and further structural details of the actinomycin-deoxyguanosine crystalline complex. J. Mol. Biol. 68, 1-20.

68. James, T.L. (1975) In 'Nuclear Magnetic Resonance in Biochemistry! Academic Press, New York.

69. Jeener, J. (1971) Ampere International Summer School Basko. Polje Yugoslavia (Unpublished lecture); Jeener, J. and Alewaelers,G. (1976) Unpublished work. Alewaeters, G. Doctoral thesis, Free University of Brussels.

70. Jeener, J. Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) Investigation of exchange processes by two dimensional NMR spectroscopy J. Chem. Phys. 71, 4546-4553.

Prog. In NMR spectroscopy 17, 1-32.

55. Gurskaya, G.V. (1965) Crystal structure of L-phenylalanine hydrochloride. Refinement and discussion of structure. Sov. Phys. Crystallogr. 9, 709-714.

56. Hare, D.R., Wemmer, D.E., Chou, S.H. and Drobny, G. (1983) Assignment of the non-exchangeable proton resonances of d(C-G-C-G-A-A-T-T-C-G-C-G) using two dimensional nuclear magnetic resonance methods. J. Mol. Biol. 171, 319-336.

57. Helene, C., Brun, F. and Yaniv, M. (1971) Fluorescence studies of interactions between Escherichia coli valyl-tRNA sysnthetase and its substrates., J. Mol. Biol. 58, 349-365; Helene, C., Montenay-Garestier, T. and Dimicoli, J.L. (1971) Interaction of tyrosine and tyramine with nucleic acids and their components. Fluorescence, nuclear magnetic resonance, and circular dichroism studies. Biochim. Biophys. Acta. 254, 349-365; Helene, C., Dimicoli, J.L. and Brun, F. (1971) Binding of tryptamine and 5-hydroxytryptamine (Serotonin) to nucleic acids. Fluorescence and proton magnetic resonance studies. Biochemistry 10, 3802-3809.

58. Helene, C. and Dimicoli, J.L. (1972) Interaction of oligopeptides containing aromatic amino acids with nucleic acids. Fluorescence and proton magnetic resonance studies. FEBS Lett. 26, 6-10.

71. Jensen, D.E. and Von Hippel, P.H. (1976) DNA "Melting" Proteins : I. Effects of Bovine pancreatic ribonuclease binding on the conformation and stability of DNA. J.Biol. Chem. 251, 7198-7214.

72. Klug, A. Jack, A., Viswamitra, M.A. Kennard,O., Shakked,Z. and Steitz, T.A. (1979) A hypothesis on a specific sequence dependent conformation of DNA and its relation to the binding of the lac repressor protein. J. Mol. Biol. 131, 669-680.

73. Kumar, A., Ernst, R.R. and Wuthrich , K. (1980) A two dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete ^{1}H-^{1}H cross-relaxation networks in biological macromolecules. Biochem. Biophys. Res. Com. 95, 1-6.

74. Kumar, A. Wagner, G., Ernst, R.R. and Wuthrich, K. (1980) Studies of J-connectivities and selective ^{1}H-^{1}H Overhauser effects in H_2O solutions of biological macromolecules by two-dimensional NMR experiments. Biochem. Biophys. Res. Com. 96, 1156-1163.

75. Kumar, N.V. and Govil, G. (1981) Role of stacking in protein—nucleic acid interactions. In 'Conformation in Biology', (Srinivasan, R. and Sarma, R.H., eds.) pp. 313-321, Adenine Press, New York.

76. Kumar, N.V. and Govil, G. (1984) Theoretical studies on protein—nucleic acid interactions. I. Interaction

energy function calculations on the interaction of Ethidium, 9-Aminoacridine, and Proflavin cations with the base-paired dinucleotides GpC and CpG. J. Amer. Chem. Soc. 101, 825-833.

101. Ogata, R.T. and Gilbert, W. (1978) An amino-terminal fragment of lac repressor binds specifically to lac operator. Proc. Natl Acad. Sci. U.S.A. 75, 5851-5854; Ogata, R.T. and Gilbert,W. (1979) DNA binding site of lac repressor probed by dimethylsulfate methylation of lac operator. J. Mol. Bio. 132, 709-728.

102. Olson, W.K. and Flory, P.J. (1972) Spatial configurations of polynucleotide chains. II. Conformational energies and average dimensions of polyribonucleotides. Biopolymers. 11, 25-56.

103. Pabo, C.O., Sauer, R.T., Sturtevant, J.M. and Ptashne, M. (1979) The λ repressor contains two domains. Proc. Natl Acad. Sci. U.S.A. 76, 1608-1612.

104. Pardi, A., Walker, R., Rapoport, H. Wider, G. and Wuthrich, K. (1983) Sequential assignments for the ^{1}H and ^{31}P atoms in the backbone of oligonucleotides by two-dimensional nuclear magnetic resonance, J. Am. Chem. Soc. 105, 1652-1653.

105. Parthasarathy, R., Ohrt, J.M. and Chheda, G.B. (1974) Conformation of N-purin-6yl carbamoyl glycine, a hypermodified base in tRNA. Biochem. Biophys. Res. Com. 57, 649-653.

of positively charged amino acids with nucleic acid fragments. Biopolymers. 23, 1979-1993.

77. Kumar, N.V. and Govil, G. (1984) Theoretical studies on protein–nucleic acid interactions. II. Hydrogen bonding of amino acid side chains with bases and base pairs of nucleic acids. Biopolymers. 23, 1995-2008.

78. Kumar, N.V. and Govil, G. (1984) Theoretical studies on protein–nucleic acid interactions III. Stacking of aromatic amino acids with bases and base pairs of nucleic acids. Biopolymers. 23, 2009-2024.

79. Kumar, M.R., Hosur, R.V., Roy, K.B., Miles, H.T. and Govil, G. (1985) Resonance assignment of the 500-MHz proton NMR spectrum of self complementary dodecanucleotide d-GGATCCGGATCC : Altered conformations at Bam HI Cleavage sites. Biochemistry. 24, 7703-7711.

80. Lakshminarayanan, A.V., and Sasisekharan, V. (1969) Stereochemistry of nucleic acids and polynucleotides. IV. Conformational energy of base-sugar Units. Biopolymers. 8, 475-488.

81. Lancelot, G. (1977) Hydrogen bonding of adenine derivatives to tyrosine side chain. Biophys. J. 17, 243-254.; Lancelot, G. (1977) Hydrogen bonding of amino acid side chains to nucleic acid bases. Biochimie. 57, 587-596.; Lancelot, G. (1977) Hydrogen bonding between nucleic

acid bases and Carboxylic acids. J. Amer. Chem. Soc. 99, 7037-7042.

82. Langlet, J., Claverie, P., Caron, F. and Boeuve , J.C. (1981) Interactions between nucleic acid bases in hydrogen bonded and stacked configurations : The role of the molecular charge distribution. Intern. J. Quantum Chem. 20, 299-338.

83. Leng, M. and Felsenfeld, G. (1966) The preferential intractions of polylysine and polyarginine with specific base sequences in DNA. Proc. Natl Acad. Sci. U.S.A. 56, 1325.

84. Loftfield, R.B., Eigner, E.A., Pastuszyn, A., Lovgren, T.N.E. and Jakubowski, H. (1980) Conformational changes during enzyme catalysis : role of water in the transition state. Proc. Natl. Acad. Sci. U.S.A. 77, 3374-3378.

85. Longworth, J.W. (1971) Luminescence of polypeptides and proteins In 'Excited states of Proteins and Nucleic Acids'. (Steiner, R.F. and Weinryb, I., eds.) pp 319-484, Plenum Press.

86. Macura, S. and Ernst, R.R. (1980) Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy. Mol. Phys. 41, 95-118.

87. Maurizot, J.C. Durand, M., Dimicoli, J.L. and Helene, C. (1973) Stud. Biophys. 40, 91. Cited in Kumar,N.V. (1983) Nucleic acid flexibility and Protein-Nucleic acid interactions. Ph.D. Thesis. TIFR, Bombay, India.

88. Maurizot, J.C., Boubault, G. and Helene, C. (1978) Interaction of aromatic residues of proteins with nucleic acids. Binding of oligopeptides to copolynucleotides of adenine and cytosine. Biochemistry. 17, 2096-2101.

81. Mayer, R., Toulme, F., Montenay-Garestier, T. and Helene, C. (1978) The role of tyrosine in the association of proteins and nucleic acids. J. Biol. Chem. 254, 75-82.

90. Mazeau, K., (Unpublished results) Centre de Biophysique Molecularie, CNRS, Orleans.

91. McKay, D.B. and Steitz, T.A. (1981) Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature. 290, 744-749.

92. McPherson, A. Jurnak, F.A., Wang, A.H.J., Molineux, I. and Rich, A. (1979) Structure at 2.3 Å resolution of the gene 5 product of bacteriophage fd : a DNA unwinding protein.J. Mol. Biol. 134, 379-400.

93. McPherson, A., Wang, A.H.J., Jurnak, F.A., Molineux, I., Kolpak, F. and Rich., A. (1980) X-ray *diffraction* studies on crystalline complexes of the gene 5 DNA

unwinding protein with deoxyoligonucleotides. J. Biol. Chem. 255, 3174-3177.

94. Montenay-Garestier, T. and Helene, C. (1971) Reflectance and luminescence studies of molecular complex formation between tryptophan and nucleic acid components in frozen aqueous solutions. Biochemistry. 10, 300-306.

95. Muller, L., Kumar, A. and Ernst, R.R. (1975) Two-dimensional Carbon-13 NMR Spectroscopy, J. Chem. Phys. 63, 5490-5491.

96. Nagayama, K., Kumar, A., Wuthrich, K. and Ernst, R.R. (1980) Experimental techniques of two-dimensional correlated spectroscopy. J. Magn . Reson. 40, 321-334.

97. Neidle, S., Berman, H.M. and Shieh, A.S. (1980) Highly structured water network in crystals of a deoxydinucleoside-drug complex. Nature. 288, 129-133.

98. Niu, C.H. and Black, S. (1979) Hydrogen bonding in solution between the uracil ring and the peptide backbone demonstrated by nuclear magnetic resonance spectroscopy. J. Biol. Chem. 254, 265-267.

99. Noggle, J.H., Schirmer, R.E. (1971) In 'The Nuclear Overhauser effect . Chemical applications'. Academic Press, New York.

100. Nuss, M.E., Marsh, F.J. and Kollman, P.A. (1979) Theoretical studies of drug-dinucleotide interactions. Empirical

87. Maurizot, J.C. Durand, M., Dimicoli, J.L. and Helene, C. (1973) Stud. Biophys. 40, 91. Cited in Kumar,N.V. (1983) Nucleic acid flexibility and Protein-Nucleic acid interactions. Ph.D. Thesis. TIFR, Bombay, India.

88. Maurizot, J.C., Boubault, G. and Helene, C. (1978) Interaction of aromatic residues of proteins with nucleic acids. Binding of oligopeptides to copolynucleotides of adenine and cytosine. Biochemistry. 17, 2096-2101.

81. Mayer, R., Toulme, F., Montenay-Garestier, T. and Helene, C. (1978) The role of tyrosine in the association of proteins and nucleic acids. J. Biol. Chem. 254, 75-82.

90. Mazeau, K., (Unpublished results) Centre de Biophysique Molecularie, CNRS, Orleans.

91. McKay, D.B. and Steitz, T.A. (1981) Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature. 290, 744-749.

92. McPherson, A. Jurnak, F.A., Wang, A.H.J., Molineux, I. and Rich, A. (1979) Structure at 2.3 Å resolution of the gene 5 product of bacteriophage fd : a DNA unwinding protein.J. Mol. Biol. 134, 379-400.

93. McPherson, A., Wang, A.H.J., Jurnak, F.A., Molineux, I., Kolpak, F. and Rich., A. (1980) X-ray diffraction studies on crystalline complexes of the gene 5 DNA

106. Parthasarathy,R., Soriano-Garcia,M. and Chheda, G.B. (1976) Bifurcated hydrogen bonds and flip-flop conformation in a modified nucleic acid base, ge Ade. Nature. 260, 807-808.

107. Pasternak, R.A. (1956) The crystal structure of Glycyl-L-Trpptophan dihydrate. Acta. Cryst. 9, 341-349.

108. Patel, D.J. (1974) Peptide antibiotic-dinucleotide interactions. Nuclear magnetic resonance investigations of complex formation between Actinomycin-D and d-pGpC in aqueous solution. Biochemistry. 13, 2388-2395; Patel,D.J.(1976) Proton and phosphorus NMR studies of d-CpG $(pCpG)_n$ duplexes in solution. Helix-coli transition and complex formation with Actinomycin-D. Biopolymers. 15, 533-558; Patel, D.J. and Canuel,L.L. (1976) Ethidium bromide-$(dC\text{-}dG\text{-}dC\text{-}dG)_2$ complex in solution : Intercalation and sequence specificity of drug binding at the tetranucleotide duplex level. Proc. Natl Acad. Sci. USA, 73, 3343-3347; Patel, D.J. and Canuel, L.L. (1977) Sequence specificity of mutagen-nucleic acid complexes in solution; Intercalation and mutagon base pair overlap geometries for proflavine binding to dC-dC-dG-dG and dG-dG-dC-dC Self-complementary duplexes. Proc. Natl Acad. Sci. USA, 74, 2624-2628.

109. Patel, D.J. (1977) d-CpCpGpG and d-GpGpCpC Self-complementary duplexes : NMR studies of the helix-coil transition Biopolymers. 16, 1635-1656.

110. Patel, D.J. and Canuel, L.L. (1979) Steroid diamine nucleic acid interactions : Partial insertion of dipyran-

106. Parthasarathy,R., Soriano-Garcia,M. and Chheda, G.B. (1976) Bifurcated hydrogen bonds and flip-flop conformation in a modified nucleic acid base, ge Ade. Nature. 260, 807-808.

107. Pasternak, R.A. (1956) The crystal structure of Glycyl-L-Trpptophan dihydrate. Acta. Cryst. 9, 341-349.

108. Patel, D.J. (1974) Peptide antibiotic-dinucleotide interactions. Nuclear magnetic resonance investigations of complex formation between Actinomycin-D and d-pGpC in aqueous solution. Biochemistry. 13, 2388-2395; Patel,D.J.(1976) Proton and phosphorus NMR studies of d-CpG $(pCpG)_n$ duplexes in solution. Helix-coli transition and complex formation with Actinomycin-D. Biopolymers. 15, 533-558; Patel, D.J. and Canuel,L.L. (1976) Ethidium bromide-$(dC\text{-}dG\text{-}dC\text{-}dG)_2$ complex in solution : Intercalation and sequence specificity of drug binding at the tetranucleotide duplex level. Proc. Natl Acad. Sci. USA, 73, 3343-3347; Patel, D.J. and Canuel, L.L. (1977) Sequence specificity of mutagen-nucleic acid complexes in solution; Intercalation and mutagon base pair overlap geometries for proflavine binding to dC-dC-dG-dG and dG-dG-dC-dC Self-complementary duplexes. Proc. Natl Acad Sci. USA, 74, 2624-2628.

109. Patel, D.J. (1977) d-CpCpGpG and d-GpGpCpC Self-complementary duplexes : NMR studies of the helix-coil transition Biopolymers. 16, 1635-1656.

110. Patel, D.J. and Canuel, L.L. (1979) Steroid diam nucleic acid interactions : Partial insertion of dipy

106. Parthasarathy,R., Soriano-Garcia,M. and Chheda, G.B. (1976) Bifurcated hydrogen bonds and flip-flop conformation in a modified nucleic acid base, ge Ade. Nature. 260, 807-808.

107. Pasternak, R.A. (1956) The crystal structure of Glycyl-L-Trpptophan dihydrate. Acta. Cryst. 9, 341-349.

108. Patel, D.J. (1974) Peptide antibiotic-dinucleotide interactions. Nuclear magnetic resonance investigations of complex formation between Actinomycin-D and d-pGpC in aqueous solution. Biochemistry. 13, 2388-2395; Patel,D.J.(1976) Proton and phosphorus NMR studies of d-CpG $(pCpG)_n$ duplexes in solution. Helix-coli transition and complex formation with Actinomycin-D. Biopolymers. 15, 533-558; Patel, D.J. and Canuel,L.L. (1976) Ethidium bromide-$(dC\text{-}dG\text{-}dC\text{-}dG)_2$ complex in solution : Intercalation and sequence specificity of drug binding at the tetranucleotide duplex level. Proc. Natl Acad. Sci. USA, 73, 3343-3347; Patel, D.J. and Canuel, L.L. (1977) Sequence specificity of mutagen-nucleic acid complexes in solution; Intercalation and mutagon base pair overlap geometries for proflavine binding to dC-dC-dG-dG and dG-dG-dC-dC Self-complementary duplexes. Proc. Natl Acad. Sci. USA, 74, 2624-2628.

109. Patel, D.J. (1977) d-CpCpGpG and d-GpGpCpC Self-complementary duplexes : NMR studies of the helix-coil transition. Biopolymers. 16, 1635-1656.

110. Patel, D.J. and Canuel, L.L. (1979) Steroid diamine nucleic acid interactions : Partial insertion of dipyran-

dium between unstacked base pairs of the poly (dA-dT) duplex in solution. Proc. Natl Acad. Sci. U.S.A. 76, 24-28.

111. Pavlovskii, A.G., Padynkova, N. Sh. and Karpeiskii, M.Ya. (1979) Structure of crystalline complexes of ribobuclease S with 8-substituted purine nucleotides. Biophysics (Moscow) 24, 327-330.

112. Perucho, M., Salas, J. and Salas, M.L. (1980) Study of the interaction of glyceraldehyde-3-phosphate dehydrogenase with DNA. Biochim. Biophys. Acta. 606, 181-195.

113. Pilz, I., Goral,K., Kratky, O., Bray,R.P., Wade-Jardetzky, N.G. and Jardetzkey, O. (1980) Small-angle X-ray studies of the quaternary structure of the lac repressor from Escherichia coli, Biochemistry. 19, 4087-4090.

114. Pople, J.A. and Segal, G.A. (1965) Approximate self consistent molecular orbital theory. I. Invariant procedures. J. Chem. Phys. 43, S129-S135.

115. Pople, J.A. and Segal, G.A. (1965) Approximate self consistent molecular orbital theory. II. Calculations with Complete Neglect of Differential Overlap. J. Chem. Phys. 43, S136-S151.

116. Porchke, D. and Gutte, B. (1981) Interaction of lac repressor fragments 33-38 (Lys-Arg-Glu-Lys-Val) with homo-oligonucleotides. FEBS Lett. 127, 63-65.

117. Porschke, D. and Jung, M. (1982) Stability decrease of RNA double helices by phenylalanine, tyrosine-and tryptophan-amides. Analysis in terms of site binding and relation to melting proteins. Nucl. Acids Res. 10, 6163-6176.

118. Pretorius, H.T., Klein, M. and Day, L.A., (1975) Gene 5 protein of fd bacteriophage. J. Biol. Chem. 250, 9262-9269.

119. Ramachandran, G.N. and Sasisekharan, V. (1968) Conformation of polypeptides and proteins Adv. Prot. Chem. 23, 283-437.

120. Reddy, B.S., Seshadri, T.P., Sakore, T.D. and Sobell, H.M. (1979) Visualization of drug-nucleic acid interactions at atomic resolution. V. Structure of two aminoacridine-dinucleoside monophosphate crystalline complexes, proflavine-5-iodocytidylyl (3'-5') guanosine and acridine orange-5-iodocytidylyl (3'-5') guanosine. J. Mol. Biol. 135, 787-812; Bhandary, K.K., Sakore, T.D. and Sobell, M. (1984) Visualization of drug-nucleic acid interactions at atomic resolution. IX. Structure of two N, N-dimethylproflavine; 5-iodocytidylyl (3'-5') guanosine Crystalline complexes. J. Biomol. Str. Dyn. 1, 1195-1217.

121. Richards, F.M. and Wyckoff, H.W. (1971) Bovine pancreatic ribonuclease. In 'The Enzymes'. 4, 647-806; Richards, F.M., Wyckoff, H.W. (1973) Ribonuclease S. In 'Atlas

of Molecular structures in Biology! (Philips, D.C. and Richards, F.M., eds.) Clarendon Press, Oxford.

122. Roberts, R.J. (1980) Restriction and modification enzymes and their recognition sequences. Nucl. Acids Res. 8, r63.

123. Saenger, W. (1984) In 'Principles of Nucleic Acid Structure' (Cantor, C.R., ed.) Springer Verlag, New York.

124. Sakore, T.D., Jain, S.C., Tsai, C.C. and Sobell, H.M. (1977) Mutagen-nucleic acid intercalative binding : Structure of a 9-amino acridine; 5-iodocytidylyl (3'-5') guanosine crystalline complex. Proc. Natl Acad. Sci. USA. 74, 188-192; Sakore, T.D., Reddy, B.S. and Sobell, H.M. (1979) Visualization of drug-nucleic acid interactions at atomic resolution. IV. structure of an amino acridine-dinucleoside monophosphate crystalline complex, 9-aminoacridine-5-iodocytidylyl (3'-5') guanosine. J. Mol. Biol. 135, 763-785.

125. Sauer , R.T., Pabo, C.O., Meyer, B.J., Ptashne, M. and Backman, K.C.(1979) Regulatory functions of the λ repressor reside in the amino-terminal domain. Nature. 279, 396-400.

126. Sauer,R.T., Yocum, R.R., Doolittle, R.F., Lewis, M. and Pabo, C.O. (1982) Homology among DNA binding proteins suggests use of a conserved super secondary structure.

Nature. 298, 447-451.

127. Scheraga, H.A. (1968) Calculations of conformations of polypeptides. Adv. Phys. Org. Chem. 6, 103-183.

128. Schulz, G.E. and Schirmer, R.H. (1979) In 'Principles of protein structure'. (Cantor, C.R., ed.) Springer-Verlag, New York.

129. Segal, D.M., Padlan, E.A., Cohen, G.H., Rudikoff, S. Potter, M. and Davies, D.R. (1974) The three dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. Proc. Natl Acad. Sci. U.S.A. 71, 4298-4302.

130. Shieh, H.S., Berman, H.M., Dabrow, M. and Neidle, S. (1980) The Structure of drug–deoxydinucleoside phosphate complex : generalized conformational behaviour of intercalation complexes with RNA and DNA fragments. Nucl. Acids Res. 8, 85-97.

131. Slater, J.C. and Kirkwood, J.G. (1931) The Van der Waals forces in gases. Phys. Rev. 37, 682-697.

132. Smith, H.O. (1979) Nucleotide sequence specificity of restriction endonucleases. Science 205, 455-462.

133. Sobell, H.M. and Jain, S.C. (1972) Stereochemistry of actinomycin binding to DNA. II. Detailed molecular model of actinomycin—DNA complex and its implications. J. Mol. Biol. 68, 21-34.

134. Sobell, H.M., Tsai, C.C., Gilbert, S.G., Jain, S.C. and Sakore, R.D. (1976) Organization of DNA in chromatin, Proc. Natl Acad. Sci. U.S.A. 73, 3068-3072.

135. Sobell, H.M., Tsai, C.C., Jain, S.C. and Gilbert, S.G. (1977) Visualization of drug-nucleic acid interactions at atomic resolution. III. Unifying structural concepts in understanding drug-DNA interactions and their broader implications in understanding protein-DNA interactions J. Mol. Biol. 114, 333-365.

136. Sobell, H.M. (1980) Structural and dynamic aspects of drug intercalation into DNA and RNA. In 'Nucleic acid geometry and Dynamics'. (Sarma, R.H., ed.) pp 289-323.

137. Sturtevant, J.M. (1977) Heat capacity and entropy changes in processes involving proteins. Proc. Natl Acad. Sci. U.S.A. 74, 2236-2240.

138. Sundaralingam, M. (1975) Principles governing nucleic acid and polynucleotide conformations. In 'Structure and conformation of Nucleic acids and Protein-Nucleic acid interactions'. (Sundaralingam, M. and Rao, S.T., eds.) pp 487-524, University Park Press, Baltimore.

139. Thomas, M.A., Ramanathan, K.V. and Kumar, A. (1983) Application of two-dimensional correlated spectroscopy to an oriented AA'BB' spin system. J. Magn. Reson. 55, 386-397.

140. Toulme, J.J., Doan, T.L. and Helene, C. (1984) Role of tryptophyl residue in the binding of gene 32 protein from phage T_4 to single-stranded DNA. Photochemical modification of tryptophan by trichloroethanol. Biochemistry. 23, 1195-1201.

141. Tsai, C.C., Jain, C.C. and Sobell, H.M. (1977), Visualization on drug-nucleic acid interactions at atomic resolution. I. Structure of an ethidium–dinucleoside monophosphate crystalline complex, ethidium; 5-iodouridylyl (3'-5') adenosine. J. Mol. Biol. 114, 301-315.; Jain, S.C., Tsai, C.C. and Sobell, H.M. (1977), Visualization of drug–nucleic acid interactions at atomic Resolution. II. Structure of an ethidium/dinucleoside monophosphate crystalline complex, ethidium; 5-iodocytidylyl (3'-5') guanosine. J. Mol. Biol. 114, 317-331.

142. Voet, D. and Bunick, G. (1978) The structures of the protein-nucleic acid interacting models, Thy $-C_3-Ind^3$ and $Ade^9-C_3-Ind^3$. Acta Cryst. A34, 572-573.

143. Voet, D. and Rich, A. (1970) The crystal structures of purines, pyrimidines and their intermolecular complexes. Prog. Nucleic Acids Res. Mol. Biol. 10, 183-265.

144. Wagner, G., Kumar, A. and Wuthrich, K. (1981) Systematic application of two-dimensional ^1H-nuclear magnetic resonance techniques for studies of proteins. II. Combined use of correlated spectroscopy and nuclear Overhauser

spectroscopy for sequential assignment of backbone resonances and elucidation of polypeptide secondary structures. Eur. J. Biochem. 114, 375-384.

145. Wodak, S.Y., Liu, M.Y. and Wyckoff, H.W. (1977) The structure of cytidilyl (2', 5') adenosine when bound to pancreatic ribonuclease S. J. Mol. Biol. 116, 855-875.

146. Wuthrich, K. (1976) In 'NMR in Biological Research : Peptides and proteins', North Holland, Amsterdam.

147. Wuthrich, K. (1986) In 'NMR of Proteins and Nucleic Acids' John Wiley, New York.

148. Yathindra, N. and Sundaralingam, M. (1973) Correlation between the backbone and side chain conformations in 5'-nucleotides. The concept of a 'rigid' nucleotide conformation. Biopolymers. 12, 297-314.