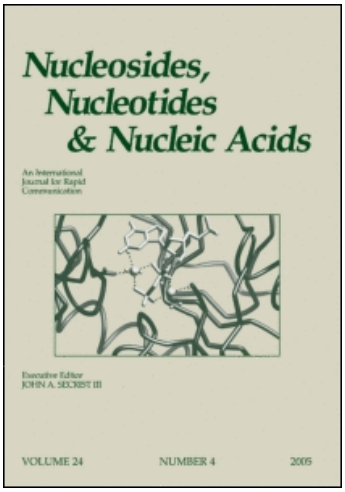

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

^{1}H NMR Structural Analysis of the Interactions of Proflavine with Self-Complementary Deoxytetranucleosides of Different Base Sequence

D. B. Davies[a]; L. N. Djimant[b]; A. N. Veselkov[b]
[a] Department of Chemistry, Birkbeck College, University of London, London, UK [b] Department of Physics, Instrument Development Institute, Sevastopol, Crimea, Ukraine

To cite this Article Davies, D. B. , Djimant, L. N. and Veselkov, A. N.(1994) '^{1}H NMR Structural Analysis of the Interactions of Proflavine with Self-Complementary Deoxytetranucleosides of Different Base Sequence', Nucleosides, Nucleotides and Nucleic Acids, 13: 1, 637 — 655
To link to this Article: DOI: 10.1080/15257779408013269
URL: http://dx.doi.org/10.1080/15257779408013269

NUCLEOSIDES & NUCLEOTIDES, 13(1-3), 637-655 (1994)

^{1}H NMR STRUCTURAL ANALYSIS OF THE INTERACTIONS OF PROFLAVINE WITH SELF-COMPLEMENTARY DEOXYTETRANUCLEOSIDES OF DIFFERENT BASE SEQUENCE†

D B Davies[1], L N Djimant[2] and A N Veselkov[2]

1. Department of Chemistry, Birkbeck College, University of London, Gordon House, 29 Gordon Square, London WC1H 0PP, UK
2. Department of Physics, Instrument Development Institute, Sevastopol 335053, Crimea, Ukraine

ABSTRACT

The molecular associations and structures of the complexes between the acridine dye, proflavine, and self-complementary deoxytetraribonucleoside triphosphates 5'-d(GpCpGpC), 5'-d(CpGpCpG), 5'-d(ApCpGpT), 5'-d(ApGpCpT) in aqueous solution have been investigated using one-dimensional and two-dimensional 500 MHz ^{1}H NMR spectroscopy.

1. INTRODUCTION

Many biologically-active aromatic molecules, such as antibiotics, carcinogens, and mutagens exert their effect primarily by interacting directly with nuclear DNA. When aromatic ligands bind to DNA they inhibit matrix synthesis which results in alterations of metabolic processes in the cell. The corresponding theories of interaction, especially the intercalation model, have made a considerable contribution to understanding the biological properties and the nature of binding of low-molecular aromatic molecules with nucleic acids. At the same time experimental and theoretical investigations of objective laws of

† This paper is dedicated to the memory of Roland K Robins and his many contributions to the chemistry of nucleosides and nucleotides

complex formation and structural features of complexes of aromatic ligands with polymer molecules are rather complicated because of the complexity of the structure of macromolecular nucleic acids and the existence of a great variety of binding sites on the polymer chain. A more realistic approach is to study the interactions of aromatic molecules with self-complementary oligonucleotides which, to a first approximation, can be considered as model systems of native nucleic acid molecules. It is known that aromatic ligands show a definite preference for binding to certain base sequences in oligonucleotide chains. Thus, the antibiotic actinomycin D preferentially binds to the GC site, but ethidium bromide and the acridine dye, proflavine, intercalate into the duplex site with pyrimidine-purine base sequence[1-7]. Such selectivity is observed in the binding of aromatic ligands with the shortest fragments of double-helical nucleic acids, the dinucleotide duplexes[7]. Investigations of the interaction of ligands with longer self-complementary oligonucleotides extend the possibilities of studying such intercalation processes and make more conclusive the transference of their properties to the polymer molecules of nucleic acids. In order to solve the problem of the selective binding of aromatic ligands, it is desirable to carry out directed experiments using specially–synthesised oligonucleotides with such preferential binding sites in different positions and various nucleotide residuals flanking the preferential site at a variety of such sites in the sequence. The shortest nucleotide sequences suitable for such investigations are tetranucleotides having three potential sites of intercalative binding.

The most complete information about the nature of the complexes of aromatic molecules with oligonucleotides is given by X-ray analysis[8]. However, the balance of forces which dominate in the crystal, particularly one of low hydration, might be expected to be very different for complex formation in solution. It should also be taken into account that the crystal structure represents a static situation whilst molecules in solution are essentially dynamic. Nuclear magnetic resonance spectroscopy is an effective experimental method to study molecular complexes in solution and the development of two-dimensional NMR techniques has opened up opportunities for investigating the mechanisms of ligand interactions with oligonucleotides in solution.

In this paper, the complex formation of a typical intercalator, proflavine, with self-complementary deoxytetraribonucleoside triphosphates 5'-d(CpGpCpG), 5'-d(GpCpGpC), 5'-d(ApCpGpT) and 5'-d(ApGpCpT) has been studied using one- and two-dimensional ^{1}H NMR spectroscopy. The first of the tetramers contains two CG sites for preferential binding of proflavine to the duplex, the second and the third tetramers have one such site in the centre, flanked from the 5'- and the 3'- ends by different nucleosides. The last tetranucleotide has no site with pyrimidine-purine base sequence.

Investigations of proflavine binding with dinucleotides have shown[7,9] that there is also some probability of intercalation of dye into purine-pyrimidine base sequences in the duplex, with the dye binding preferentially to such sites in the monomeric nucleotide chain. Data from homonuclear correlation 2D-NMR spectroscopy (2D-COSY and 2D-NOESY) were used to assign proton resonances in the ^1H NMR spectra of mixed solutions and to determine qualitatively the binding sites of the dye with tetranucleotides. The analysis of the interaction of proflavine with these tetranucleotides in solution, the influence of the base sequence of the nucleotide chain on the nature of complex formation, and the structural features of complexes of different types have been made using experimental concentration dependences of chemical shifts of non-exchangeable protons of nucleotides and dye.

2. MATERIALS AND METHODS

Methods for the preparation of solutions and performance of experimental procedures are described in the previous paper[10]. Deoxytetranucleotides with defined base sequences were synthesised by OSWEL DNA SERVICE (Edinburgh University, UK). The tetranucleotides and the acridine dye proflavine ('Sigma') were lyophilised three times from 99,95% D_2O then dissolved in deuteriated 0.1 M phosphate buffer, pD = 6.6. 500 MHz ^1H NMR spectra were recorded on a JEOL GSX-500 spectrometer with the residual water peak saturated during relaxation. The complete assignments of proton signals had been made previously using 2D-COSY and 2D-NOESY spectra of tetranucleotides[10], enabling straightforward assignment of the spectra of mixed solutions. 2D-NOE spectra of solutions of tetranucleotides with dye were used to elucidate the possible through-space connectivities between the protons of these molecules.

3. RESULTS AND DISCUSSION

In principle, the interpretation of 2D-NOESY spectra provides information about internuclear contacts between interacting molecules and hence about the conformation of oligonucleotides in a complex with dye in mixed solutions. The merit of NOE is that it enables direct determination of interproton distances from NOE build-up rates. Reid *et al* [11] have shown that qualitative interpretation of 2D-NOESY cross-peaks can be made when NOESY experiments are carried out at mixing times not exceeding 100 ms. Only in this case can one consider the effective correlation time of interproton vectors to be the same for all interacting nuclei and calculate internuclear distances using the corresponding equations for the cross-relaxation rate of two-spin systems. Under experimental conditions it is not always possible to reach the full intensities of cross-peaks between protons of the intercalating ligand and the nucleotides which is required for reliable determination of corresponding distances and the structures of the complexes. However, even in such

cases, 2D-NOE spectra enable important information to be determined about possible binding sites of the ligand with the oligomer and conclusions about molecular conformational states can be made.

An example is given by the 2D-NOE spectra of proflavine solutions with 5'-d(CpGpCpG) presented in Fig. 1. Together with cross-peaks defining intramolecular interactions between nuclei (connectivities between different protons are shown by dotted lines), cross-peaks are observed between some dye protons and nucleosides (filled lines) for proflavine with 5'-d(CpGpCpG); cross-peaks of approximately equal intensities can be observed between protons H_X-H5(C1) and H_X-H5(C3) as well as between H_B-H2'(G2) and H_B-H2'(G4) protons. The existence of such internuclear spin-spin interactions supports the assumption that the dye intercalates into both CG sites of the tetramer with approximately the same probability. In the spectrum of the solution of the dye with 5'-d(GpCpGpC) we have observed cross-peaks between protons of proflavine and internal nucleosides (H_B-H2'(G3), H_B-H8(G3), and H_X-H5(C2)) - the latter confirming preferential binding of ligand to the central CG site of the tetramer. A similar conclusion about dye intercalation into the CG-site of the tetranucleotide 5'-d(ApCpGpT) can be drawn from the existence of 2D-NOE cross-peaks between H_B-H2'(G3) and H_X-H5(C2). Although there is no site with pyrimidine-purine base sequence in the tetramer 5'-d(ApGpCpT), the spectra of mixed solutions with dye also provide evidence about preferential binding of the dye to the GC site in the centre of this sequence. It should be noted that the intensites of all the cross-peaks due to intermolecular interactions are rather low compared to those observed for intramolecular interactions which may be due to the fact that distances are relatively high between protons of ligand and tetranucleotides at the intercalation site. In addition, the number of intermolecular cross-peaks in all the 2D-NOE spectra is rather small which may be of significance not only for proflavine but also for other intercalating dyes containing planar chromophores without large side groups or chains. As a result, the determination of the geometry of the complexes of such ligands with oligonucleotides cannot be solved using only 2D-NOESY data.

Additional geometrical information can be obtained from a knowledge of the limiting values of proton chemical shifts of the molecules in the bound state, enabling structural characteristics of the complexes to be determined using quantum mechanical calculations of shielding effects of aromatic molecules[12]. For solution studies, especially in the case of short oligonucleotides, a rather complex dynamic equilibrium of the interacting molecules takes place; an equilibrium which includes monomers and dimers of both the tetranucleotide and dye and their complexes of different types. Experimentally-observed

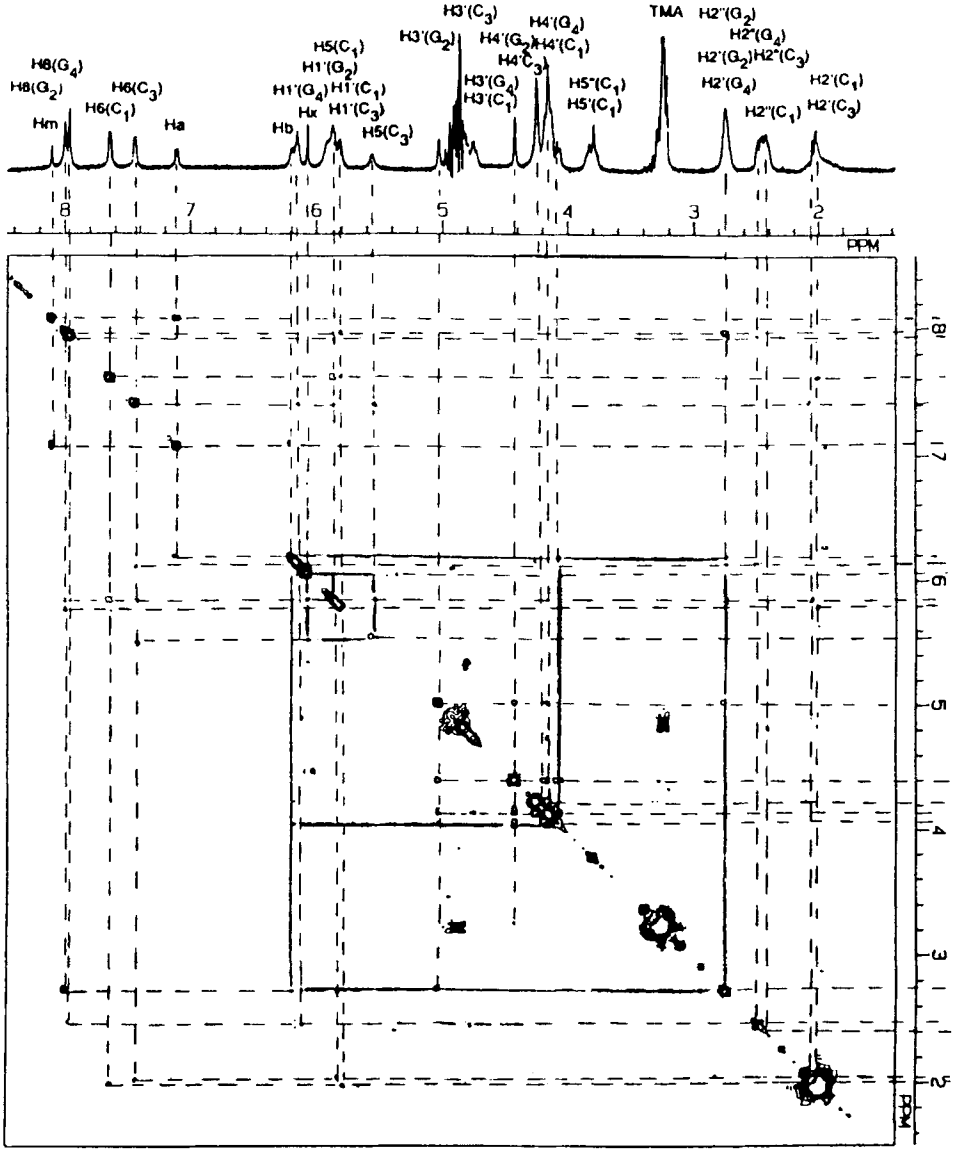

FIG.1. 2D-NOE spectrum of the mixture of proflavine and tetranucleotide 5'-d(CpGpCpG) at initial concentrations of proflavine P_0=0.76 mM and tetranucleotide N_0=1.2mM (T=298 K).

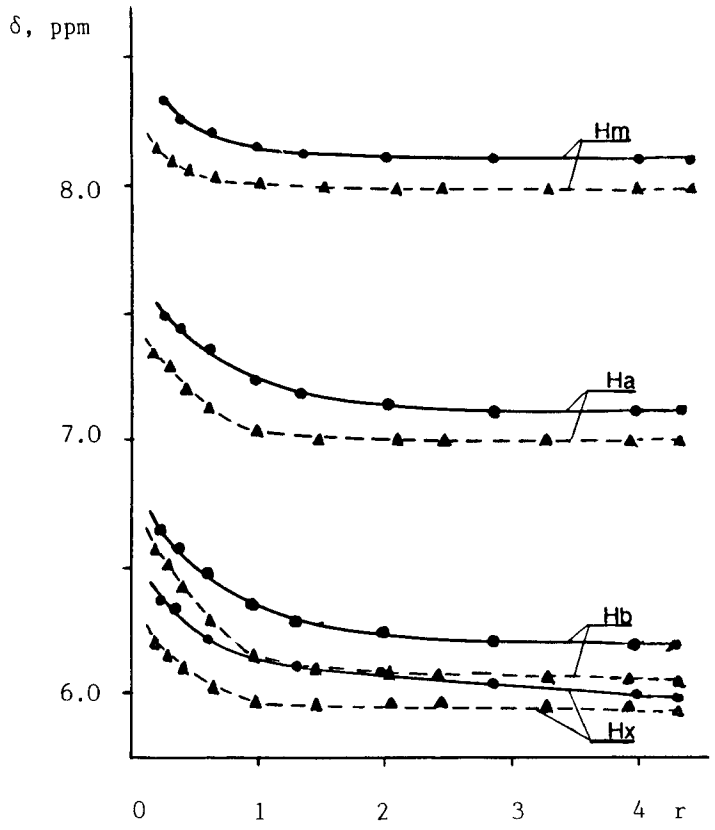

FIG. 2. Dependence of proflavine proton chemical shifts on r, the ratio of initial dye and double-stranded tetranucleotide concentrations ($r = N_O/P_O$) at constant proflavine concentration in solution. The solid lines represent the mixture of proflavine and 5'-d(GpCpGpC) at $P_O = 0.546$ mM and the dashed lines represent the mixture of proflavine and 5'–d(CpGpCpG), at $P_O = 0.76$ mM.

proton chemical shifts of the molecules are the result of a sum of contributions of monomer and associated forms, fractions of which depend on the initial concentrations of molecules in solution. Measurements of the concentration dependence of proton chemical shifts provides information to determine the types of complexes being formed, the equilibrium reaction constants and the limiting values of chemical shifts of bound molecules[7,9,13,14]. For example, the concentration dependences of chemical shifts, δ, of proflavine protons in solution with the tetranucleotides 5'-d(GpCpGpC) and 5'-d(CpGpCpG) are shown in Fig. 2. The concentration parameter, r ($=N_o/P_o$), is the ratio of initial concentrations of

tetranucleotide duplex and dye. It is seen that pronounced changes of dye proton chemical shifts take place at small r values when the dye content in solution is relatively high whereas, when $r > 1$, the tetranucleotide is in abundance in solution and the probability of proflavine binding is high enough so that the concentration dependence for δ is less pronounced. Similar curves of $\delta=f(r)$ are observed for solutions of dye with 5'-d(ApCpGpT) and 5'-d(ApGpCpT).

3.1 Proflavine interaction with 5'-d(CGCG) and 5'-d(GCGC)

In order to estimate quantitatively the complexation of proflavine with tetranucleotides, different schemes of reactions in solution have been analysed. In the case of the interaction of dye with isomeric sequences, 5'-d(GpCpGpC) and 5'-d(CpGpCpG), the best agreement with experimental data is given by the following scheme for molecular association:

$$\begin{array}{ll} \overset{K_P}{P + P \Leftrightarrow P_2} & \overset{K_T}{N + N \Leftrightarrow N_2} \\ \overset{K_1}{P + N \Leftrightarrow PN} & \overset{K_2}{PN + P \Leftrightarrow P_2N} \\ \overset{K_3}{P + N_2 \Leftrightarrow PN_2} & \overset{K_4}{PN_2 + P \Leftrightarrow P_2N_2} \end{array} \qquad (1)$$

In this scheme of complex formation, an assumption has been made about the basic role of double-component interactions in solution. Taking into consideration the mass law equations for reactions (1) and the mass conservation law, we obtain a system of equations:

$$\begin{aligned} &2(K_P + K_1K_2N + K_TK_3K_4N^2)P^2 + (1 + K_1N + K_TK_3N^2)P - P_0 = 0 \\ &2(K_T + K_TK_3P + K_TK_3K_4P^2)N^2 + (1 + K_1P + K_1K_2P^2)N - N_0 = 0 \end{aligned} \qquad (2)$$

where P and N are equilibrium molar concentrations of proflavine and tetranucleotide monomers, respectively.

Chemical shifts were calculated using an additive model:

$$\delta = (P/P_0)(\delta_m + 2K_P P\delta_d + K_1N\delta_1 + 2K_1K_2PN\delta_2 + K_TK_3N^2\delta_3 + 2K_TK_3K_4PN^2\delta_4) \qquad (3)$$

In this relation δ_m, δ_d, δ_1, δ_2, δ_3, and δ_4 are the proton chemical shifts of the dye molecule in the monomer, dimer, and 1:1 (PN), 2:1 (P_2N), 1:2 (PN_2), and 2:2 (P_2N_2) complexes, respectively. The values of δ_m and δ_d were determined previously[9] from investigations of proflavine self-association using concentration dependences of dye proton chemical shifts in the same solvent system (0.1 M phosphate buffer). The value of proflavine dimerisation constant, K_P, averages 300 - 350 M^{-1} at temperatures T = 293 - 298 K[15]. Values of the constant of self-association, K_T, of tetranucleotides 5'-d(GpCpGpC) and 5'-d(CpGpCpG), were determined independently from analysis of the concentration dependencies of ^1H NMR spectra of tetramers[10] and, under the conditions of this experiment, K_T =1000(±120) (mol of duplex)$^{-1}$ and K_T =1300(±160) (mol of duplex)$^{-1}$, respectively. It follows that the observed chemical shift, δ, is a function of the unknown quantities $\delta_1 - \delta_4$, $K_1 - K_4$, P and N. As the concentrations of dye, P, and tetranucleotide, N, in turn, depend on K_1 - K_4, P and N can be found from the solution of the system of equations (2). Parameters δ_1, δ_2, δ_3, δ_4, K_1, K_2, K_3, and K_4 of equation (3) were determined using the variational method of data analysis by minimization of the quadratic discrepancy function[9] using experimental concentration dependences of chemical shifts for the different dye protons. Such a multiparametric problem required solution in two stages: the initial approximation was found by the method of accidental search of an unconditioned minimum- and further optimisation was carried out by the modified simplex method of Nelder and Mead[16,17]. A detailed discussion of the computational procedure and the evolution of different models used in the analysis of the concentration dependences of proton chemical shifts has been published[18] for one of the systems studied, proflavine with 5'-d(GCGC). The analysis begins with the simplest case when the only reactions taken into account are the self-association of proflavine and the tetranucleotide, and 1:1 and 1:2 proflavine–tetranucleotide complex formation; calculations are then made for the sequential inclusion of new reactions in scheme (1). The parameters obtained at each sequential step serve as the first approximation for the calculation of parameters in the more complicated scheme. The actual fit of the curves (discrepancy between experimental and calculated points) for the general model in scheme (1) did not exceed 1%.

The calculated values of equilibrium constants of complex formation K_i, and induced high field chemical shifts for the protons of proflavine, $\Delta\delta_i = \delta_m - \delta_i$, where i =1,2,3,4 for the corresponding type of complex, are presented in Table 1. The values of $\Delta\delta_i$ define the shielding effects of nucleosides on dye protons in the complex. Constants K_1 and K_2, of 1:1 and 2:1 complex formation of proflavine with the single strand of 5'-d(GpCpGpC) tetranucleotide are approximately equal in value due to the existence in this sequence of two GC sites of preferential binding of proflavine with monomer[9]. The tetranucleotide 5'-

TABLE 1: Calculated values of parameters of complex formation of proflavine with 5'-d(GpCpGpC) and 5'-d(CpGpCpG)

Tetranucleotide	5'-d(GpCpGpC)				5'-d(CpGpCpG)			
Parameter	H_M	H_A	H_B	H_X	H_M	H_A	H_B	H_X
$\Delta\delta_1$/ppm	1.02	0.89	1.01	0.95	1.03	1.01	1.28	0.91
$\Delta\delta_2$/ppm	0.95	0.94	1.06	0.90	1.30	1.12	1.21	1.20
$\Delta\delta_3$/ppm	0.49	0.73	0.91	0.88	0.75	0.98	0.96	0.93
$\Delta\delta_4$/ppm	0.55	0.75	1.10	0.62	0.59	0.84	0.98	0.66
K_1/M^{-1}	3800 ± 600				6200 ± 800			
K_2/M^{-1}	2500 ± 400				3630 ± 660			
K_3/M^{-1}	9400 ± 1000				17600 ± 1600			
K_4/M^{-1}	520 ± 120				10900 ± 900			

d(CpGpCpG) contains one GC site flanked from both sides by nucleotides which increases the probability of vertical stacking interactions between guanine and cytosine bases in the centre of the sequence and facilitates binding of the planar dye molecule with the tetramer; this is the likely reason that the constant of complex formation K_1, of proflavine with d(CpGpCpG), is higher than the corresponding constant for the isomeric sequence.

Binding of the second dye molecule with 5'-d(CpGpCpG) monomer, as seen by comparison of the constants K_2 and K_1, takes place with somewhat less probability than the first. If one dye molecule preferentially binds to the central GC site, then the process of binding of the other molecule with adjacent sites of such sequence will be, obviously, nonco-operative. The existence of preferential dye binding to pyrimidine-purine base sequences in the duplex is supported by the calculated values of the constants K_3 and K_4, which are substantially different for the isomeric tetranucleotides considered. The equilibrium constant, K_3, of 1:2 complex formation of dye with 5'-d(CpGpCpG), a sequence containing two CG sites, is approximately double that for its isomer with one such site which indicates that proflavine intercalates into each CG site of these double

helical sequences with approximately the same probability. This conclusion is supported by the disposition and relative intensities of cross-peaks between protons of proflavine and nucleotides in 2D-NOE spectra. Comparison of the equilibrium association constants, K_4 and K_3, shows that the binding of the second molecule with 5'-(GpCpGpC) is highly nonco-operative providing additional confirmation that the dye molecule preferentially intercalates into CG sites so that the binding of the second molecule to such a duplex is very difficult. On the other hand, in the interaction of proflavine with 5'-d(CpGpCpG), where there are two sites with pyrimidine-purine base sequences, the reaction constant, K_4, is high, approximately 20 times greater than for the isomeric tetranucleotide. When taken together, these results confirm that the formation of 2:2 complexes of proflavine with tetranucleotide duplexes correspond to the 'excluded neighbour' model[19] according to which intercalation of the dye molecule between adjacent base pairs is impossible.

Relatively large values of induced chemical shifts, $\Delta\delta_i$ (Table 1), indicate the considerable shielding effects of nucleosides on all the proflavine protons which results from intercalation of the dye chromophore between the base rings of the tetranucleotides. Quantum mechanical calculation of the shielding effects of nucleic acid bases[12] predicts maximum-induced chemical shifts around 1.7 - 1.8 ppm for the protons of a planar molecule situated between guanine and cytosine at a distance of 3.4 Å from the base planes. The calculated values of $\Delta\delta_i$ for all dye protons in the complexes are in the range 0.5 - 1.3 ppm indicating a rather extensive overlap between the planes of the interacting molecules. It is significant that the values of induced chemical shifts $\Delta\delta_1$ and $\Delta\delta_2$ for the complexes of dye with the monomers of these tetranucleotides are somewhat higher than the values of $\Delta\delta_3$ and $\Delta\delta_4$ for the complexes with the duplex (Table 1), despite the fact that the shielding of dye protons results from nucleosides of only a single-stranded tetramer. Preliminary calculations show that in such cases the necessary shielding of all the proflavine protons can be achieved only at distances less than 3.4 Å between the ligand chromophore and adjacent tetranucleotide bases, which is made possible by the high conformational freedom of the single-strand nucleotide sequence. As for the complexes of dye with the double helix, comparison of $\Delta\delta_3$ and $\Delta\delta_4$ values shows that the structures of 1:2 and 2:2 complexes are not substantially different for both tetranucleotides studied although a somewhat higher shielding of H_M and H_A dye protons in the 1:2 complex of proflavine with 5'-d(CpGpCpG) compared with 5'-d(GpCpGpC) indicates deeper intercalation of ligand into CG sites of the double helix.

3.2 Proflavine interaction with 5'-d(ApCpGpT) and 5'-d(ApGpCpT)

The corresponding analysis of concentration dependences of dye proton chemical shifts has also been carried out for solution mixtures with the

tetramers 5'-d(ApCpGpT) and 5'-d(ApGpCpT), in which central CG and GC sites are flanked by adenine and thymine. However, calculations using scheme (1) for these sequences did not give satisfactory agreement with the experiment unless the possibility of 1:2 complex formation of proflavine with tetranucleotide (PN_2) is considered in two different ways: direct binding of dye with the duplex and formation of this complex by interaction of the tetranucleotide monomer with 1:1 complex PN, where proflavine acts as a 'nucleation centre'.

$$\begin{array}{lll} \overset{K_P}{P + P \Leftrightarrow P_2} & & \overset{K_T}{N + N \Leftrightarrow N_2} \\ \overset{K_1}{P + N \Leftrightarrow PN} & & \overset{K_2}{P + PN \Leftrightarrow P_2N} \qquad (4) \\ \overset{K_3}{PN + N \Leftrightarrow PN_2} \quad (a) & & \overset{K_4}{P + N_2 \Leftrightarrow PN_2} \quad (b) \end{array}$$

It is necessary to consider the structure of the complex by means of the 'nucleation centre' (Reaction 4a) because the equilibrium self-association constants of the tetramers with AT base pairs at the ends of the sequences are substantially lower than K_T for tetranucleotides containing only CG base pairs. The values of self-association constants turned out to be approximately equal for 5'-d(ApCpGpT) and 5'-d(ApGpCpT) and averaged 160(±30) (mol of duplex)$^{-1}$ at T = 293 K[10]. The reaction of 2:2 complex formation of dye with these tetramers, that contain only one CG or GC site, may be neglected in a general equilibrium in solution as shown by corresponding calculations in this work.

According to the reactions scheme (4), the system of equations determining the equilibrium molar concentrations of interacting monomers, P and N, is:

$$\begin{aligned} &2(K_P + K_1K_2N)P^2 + (1 + K_1N + K_1K_3N^2 + K_TK_4N^2)P - P_0 = 0 \\ &2(K_T + K_1K_3P + K_TK_4P)N^2 + (1 + K_1P + K_1K_2P^2)N - N_0 = 0 \end{aligned} \qquad (5)$$

and an expression for the chemical shifts of proflavine protons can be written in the form:

$$\delta = (P/P_0)(\delta_m + 2K_PP\delta_d + K_1N\delta_1 + 2K_1K_2PN\delta_2 + K_1K_3N^2\delta_3 + K_TK_4N^2\delta_4) \qquad (6)$$

Using equation (6), the calculated values of the parameters of complex formation of proflavine with 5'-d(ApCpGpT) and 5'-d(ApGpCpT) are presented in Table 2. For both

TABLE 2: Calculated values of parameters of complex formation of proflavine with 5'd–(ApCpGpT) and 5'–d(ApGpCpT).

Tetranucleotide	5'-d(ApCpGpT)				5'-d(ApGpCpT)			
Parameter	H_M	H_A	H_B	H_X	H_M	H_A	H_B	H_X
$\Delta\delta_1$/ppm	1.03	0.80	0.72	1.01	1.00	0.89	0.86	1.18
$\Delta\delta_2$/ppm	0.53	0.60	0.45	0.95	0.82	0.68	0.79	0.70
$\Delta\delta_3$/ppm	0.83	0.87	0.87	0.93	0.81	0.89	0.87	0.79
$\Delta\delta_4$/ppm	0.79	0.81	0.91	0.96	0.76	0.75	1.04	0.84
K_1/M^{-1}	22000 ± 1800				24300 ± 1500			
K_2/M^{-1}	1520 ± 380				470 ± 150			
K_3/M^{-1}	15700 ± 2300				3400 ± 4500			
K_4/M^{-1}	4880 ± 1220				7000 ± 1200			

tetranucleotides it is significant that the equilibrium reaction constants K_2 are much smaller than K_1, so that the binding of the second proflavine molecule with a single strand is highly nonco-operative. Similarly, intercalation of the second molecule into the duplex forming a 2:2 complex is also nonco-operative as our additional calculations have shown. It is likely that dye molecules preferentially intercalate into the central site of the tetramers so that the binding of one more proflavine molecule either to the single strand of tetranucleotide or to the duplex is improbable.

It is noteworthy that the equilibrium constants of 1:1 complex formation of proflavine with these tetranucleotides are much higher than those for the corresponding complexes of this dye with dinucleotides CpG and GpC which average $K_1 = 140(\pm 130)\ M^{-1}$ and $K_1 = 750(\pm 150)\ M^{-1}$, respectively[9]. The substantial difference in complex formation of proflavine with monomers of dinucleotides and tetramers which have central sites with the same nucleotide sequences is probably due to the different intramolecular stacking between adjacent bases in dinucleotides and tetranucleotides. In the latter, the supposed binding site with dye is flanked by nucleosides from the 5'- and the 3'- ends of the sequence and, in

this connection, there is less probability, compared with dinucleotides, of 'opening' of the base planes in the central site of the tetramer to provide more favourable conditions for the binding of the planar dye molecule.

The constant, K_3, defining 1:2 complex formation through a 'nucleation centre' (reaction 4a), turned out to be much higher for the 5'-d(ApCpGpT) sequence than for 5'-d(ApGpCpT). This can be explained by taking into account the results of the investigation of the complex formation of proflavine with dinucleotides CpG and GpC where it was shown[9] that, when the dye interacts with single strands of dinucleotides, the structures of the 1:1 complexes with CpG and GpC are substantially different. In the 1:1 complex of proflavine with CpG, the dye molecule is orientated with respect to the tetranucleotide strand by the side aromatic ring of the chromophore[9]. Such an orientation of the planar proflavine molecule in the CG site of the tetramer creates favourable conditions for the binding of the second single strand tetranucleotide molecule through the 'nucleation centre' (reaction 4a) forming the 1:2 complex and it is expected that the value of the constant K_3, will be high. In the 1:1 complex of proflavine with GpC, the dye molecule is orientated with respect to the nucleotide chain by the central aromatic ring of the chromophore so that considerable overlap of the planes of the dye molecule and bases takes place[9]. Additional stabilisation in such a structure is possible due to electrostatic interactions of the positively charged nitrogen atom of proflavine with the negative phosphate of the dinucleotide. In this case, the arrangement of the planar proflavine molecule in the GC site of the tetramer is not optimum for binding the second single strand tetranucleotide molecule through the 'nucleation centre', because it is not provided with sufficient stacking interactions between the dye chromophore and the bases of the binding monomer of the tetranucleotide. It is likely that the formation of such a complex should be accompanied by reorientation of the dye molecule at the intercalation site with additional expenditure of energy resulting in a lower K. As for the possibility of direct intercalation of the proflavine molecule into the duplex, a comparison of K_4 values shows that they are approximately equal for both the tetranucleotides studied.

It follows from the data presented in Table 2 that the calculated values of induced chemical shifts of proflavine protons $\Delta\delta_3$ and $\Delta\delta_4$ in the 1:2 complexes are close to each other for both tetranucletides which confirms that the geometry of the 1:2 complex does not depend on the mechanism of its formation. The values of $\Delta\delta_1$ and $\Delta\delta_2$ differ significantly between each proton in proflavine and the induced chemical shifts for all the proflavine protons are considerably smaller in the 2:1 compared with the 1:1 complex. It should be noted that the calculated value of $\Delta\delta_2$ in the 2:1 complex is the mean value for the protons

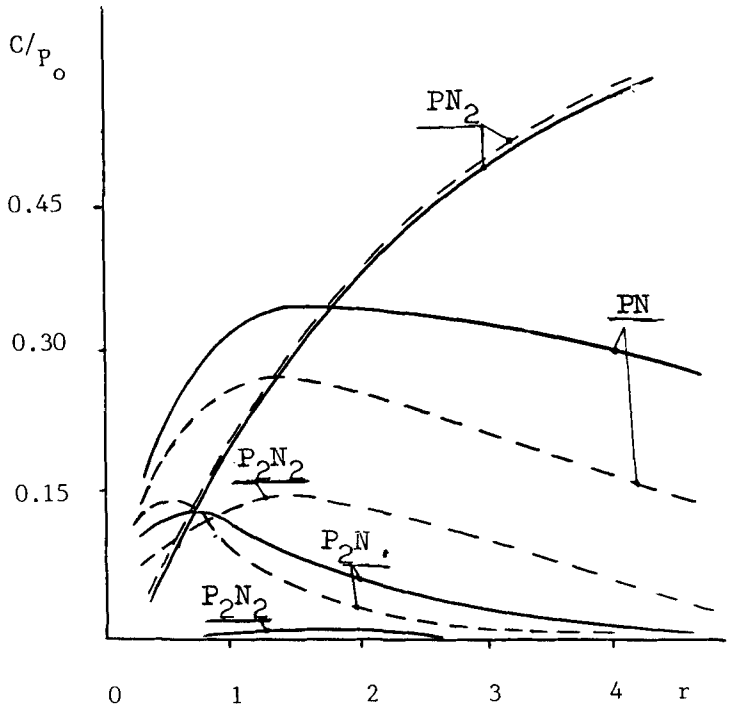

FIG. 3: Calculated relative content of molecular complexes in solution as a function of r = N_o/P_o. The solid lines represent the mixture of proflavine and 5'-d(GpCpGpC) and the dashed lines represent the mixture of proflavine and 5'-d(CpGpCpG).

of the two dye molecules bound with the tetranucleotide. Supposing that one dye molecule intercalates into the central site, then the other one binds, in all probability from the outside, by means of vertical stacking interactions to the end nucleotides because shielding of its protons is insignificant.

3.3 Concentration dependence of molecular complex formation

The relative content of molecular complexes in solution as a function of r (=N_o/P_o, the ratio of the initial concentrations of tetranucleotide duplex and dye) has been determined using the calculated values of reaction constants in this work. The contribution to the general equilibrium in solution of different types of complexes is determined not only by the values of equilibrium reaction constants, but also by the nature of nucleotide sequence and the value of r. With increasing r, a monotonic growth of the fraction of 1:2 complexes, PN_2, is observed for all the tetranucleotides. For tetramers 5'-d(GpCpGpC) and 5'-d(CpGpCpG), this is due to the high probability of formation of such structures with

increasing duplex concentration in solution (Fig.3). When proflavine interacts with tetranucleotides 5'-d(ApCpGpT) and 5'-d(ApGpCpT), 1:2 complexes at relatively high r values are formed mainly through the 'nucleation centre' due to the considerable concentration of tetranucleotide monomers and 1:1 complexes in solution. The contribution of the 2:2 complex (Fig. 3, P_2N_2) is substantial only for solutions of dye with tetramer 5'-d(CpGpCpG) that contains two CG sites unlike any other sequence studied; the decrease of the fraction of this complex at large r is due to the small concentration of free dye in solution. It is noteworthy that the relative amount of 2:1 complexes, P_2N, of dye with 5'-d(ApCpGpT) and 5'-d(ApGpCpT) is small and becomes negligible already at $r \approx 1$, providing evidence of the existence of only one dye binding site in such sequences. At the same time, the fraction of 2:1 complex is much higher for the other tetramers containing two potential binding sites (Fig.3). It is significant that some curves are not monotonic and have maximums at r values corresponding, as a rule, to stoichiometric relations of tetranucleotide and dye concentrations for complex formation.

The analysis of the complex equilibrium in solution shows that the contribution of each type of complex to the experimentally-observed proton chemical shifts is determined by the magnitudes of equilibrium reaction constants and the values of initial concentrations of interacting molecules and depends substantially on the base sequence in the nucleotide chain. In this connection, a purely qualitative interpretation of observed NMR data without the corresponding quantitative analysis of the system (see, for example, references 20-21) does not allow, in our opinion, the equilibrium content of different associations in solution to be determined properly, as well as the limiting proton chemical shifts of bound ligands which may lead to substantial errors in determining the distinctive structural features of the complexes being formed.

3.4 Structures of the complexes

In this paper, the structures of the complexes of proflavine with tetranucleotides have been determined by a combination of 2D-NOE spectra and calculated induced chemical shifts of dye protons. Assuming that the planes of the base pairs in the complex are separated by 6.8 Å, the most favourable structures of the 1:2 complexes of proflavine with the four tetranucleotide duplexes were found using isoshielding maps computed by Giessner-Prettre and Pullman[12]. The results of these calculations are shown in Fig. 4 for two representative tetranucleotides, 5'-d(CpGpCpG) and 5'-d(ApGpCpT). In these structures, proflavine intercalates into the CG site of the 5'-d(CpGpCpG) sequence and the GC site of the tetranucleotide 5'-d(ApGpCpT) from the side of N(7) and N(4) of the base pairs, i.e. from the major groove of the minihelix. The dye chromophore is parallel to the

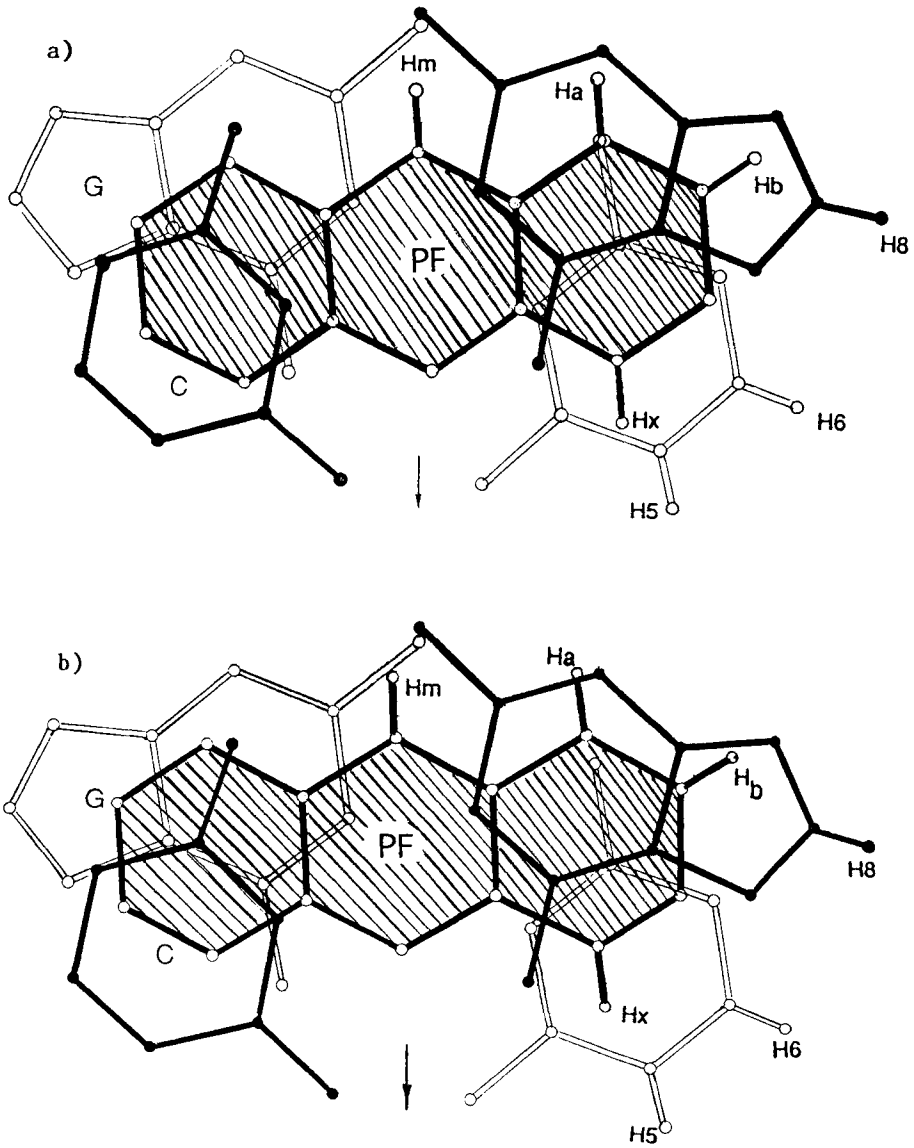

FIG. 4. Calculated structures of 1:2 complexes of proflavine with tetranucleotides: a) proflavine with d(CpGpCpG), b) proflavine with d(ApGpCpT). The structures are viewed perpendicular to the planes of the base pairs and dye chromophore. The atoms of the uppermost base pairs, bonds between them, and the bonds of intercalated proflavine are in black. The dye chromophore is shaded. An arrow shows the direction to major groove of the helix.

base planes and is situated at equal distances of 3.4 Å from them. Intercalation of the dye molecule into these tetramer sites is corroborated by the existence of cross-peaks between proflavine protons and the corresponding nucleosides in 2D-NOE spectra. In determining the structure of the complex, different geometries of the base pairs[23] at the intercalation site for the helix in the B–form have been considered. The ring current effect of the nucleotide bases provides the main contribution to the shielding effect of neighbouring molecules.[12] As the ring current effect decreases rapidly with distance (with a $1/r^3$ dependence), only the influence of adjacent base pairs in the double helix should be taken into consideration for calculations of the structures of the complexes of dye with the tetranucleotide duplex. Computational analysis of the ring current effects of adjacent base pairs on proflavin chemical shifts shows that, within the limits allowed for double-helical DNA in solution, the variation of angular parameters κ (buckle), ρ (roll) and ω (propeller twist) have little influence on the induced chemical shifts of the symmetrical molecule, proflavine, and can be neglected compared to the influence of the translational parameters D_x(shift), D_y(slide) and D_z (rise) and angular parameters τ (tilt) and Ω (twist) of the base pairs; only the effect of variation of these latter five parameters was tested in the calculation of the intercalculated structure from limiting proton chemical shifts. The notations and nomenclature of translational and angular parameters of the helical structures are taken from reference 24.

It is significant that the structures of all the 1:2 complexes formed by intercalation of proflavine into the CG sites of all the tetranucleotides are not substantially different from each other. These structres are in good agreement with that for the 1:2 complex of proflavine with d(CpG) determined by crystal structure analysis[8]. Similar, but slightly different geometry is obtained when the dye intercalates into the GC site of tetranucleotide 5'-d(ApGpCpT) although the differences in the structures are not large and the general orientation of the chromophore relative to the base pairs is preserved (Fig.4). Such similarity of the structures of the 2:1 complexes is also observed when proflavine binds to isomeric CG and GC duplexes[7,9]. Hence, it can be concluded that, for 1:2 complexes, the mutual orientations of the planes of the base pairs and the dye chromophore do not depend on the base sequence in the CG sites of the tetramers nor on the type of nucleoside flanking the ligand binding site. At the same time, as has been shown above, the nature of nucleotide sequences has a substantial influence on the values of equilibrium association constants and hence on the probability of formation of such structures in solution. It can be assumed that selectivity of proflavine binding with the pyrimidine-purine base sequence is due mainly to different conformational transformations of tetranucleotides during dye intercalation. Investigations of intercalated complexes of dye with dinucleotides in crystals[25] confirm that the structure of sugar-phosphate backbone in such associations depends on the nucleotide sequence.

Determination of common features in the selectivity and specificity of the interaction of biologically-active aromatic molecules with nucleic acids is a very important problem. Analysis of the structures of crystal complexes of aromatic ligands with oligonucleotides summarised the most essential physico-chemical factors that determine intercalative recognition[26]. They are the following: conformational states of different nucleotide sequences sensitive to both the planar chromophore of an incoming ligand and its side chains; steric differences between major and minor duplex grooves that can play an important role when an intercalant has a bulky side-chain or grouping attached; the existence at the ligand of atom groups with potential hydrogen-bonding ability with donors or acceptors in the bases of nucleic acids; the existence of cationic charge on either the ligand chromophore itself or its side chains. One might expect that these factors will also contribute the specificity of binding of aromatic molecules with oligonucleotides in aqueous solution.

ACKNOWLEDGEMENT

We thank the British Council for a Senior Travel Fellowship (to A.N.V) and the Univervity of London Intercollegiate Research Service for access to the 500 MHz Biomedical NMR facility (at Birkbeck College).

REFERENCES

1. Krugh T R, Reinhardt C G, *J.Mol. Biol.*, **1975**, *97*, pp 133 - 162
2. Patel D J, Canuel L L *Proc. Natl. Acad. Sci. USA*, **1977**, *74*, pp 2624 - 2628
3. Kastrup R V, Young M A, Krugh T R *Biochemistry*, **1978**, *17*, N° 23, pp 4855 - 4864
4. Scott E V, Jones R L, Banville D L, Zon G, Marzilli L G, Wilson W D, *Biochemistry*, **1988**, *27*, pp 915 - 923
5. Delepierre M, Heijenoort C V, Igolen J, Pothier J, Le Bret M, Roques B P *J. Biomol. Struct. Dyn.*, **1987**, *7*, N° 3, pp 557 - 589
6. Chen F-M *Biochemistry*, **1988**, *27*, pp 6393 - 6397
7. Veselkov A N, Djimant L N, Baranovsky S F, *Molecular Biol.* (in Russian), **1986**, *20*, N° 5, pp 1244 - 1250
8. Neidle S, Berman H, *Prog. Biophys. Molec. Biol.*, **1983**, *41*, pp 43 - 66
9. Veselkov A N, Karawajew L S, Djimant L N *Studia Biophys.*, **1987**, *120*, N° 1, pp 87 - 107
10. Veselkov A N, Davies D B, Djimant L N, Parkes H G, Shipp D *Biopolymers and cell* (in Russian), **1991**, *7*, N° 5, pp 234 - 240

11. Reid B R, Banks K, Flynn P, Nerdal W *Biochemistry*, **1989**, *28*, pp 10001 - 10007

12. Giessner-Prettre C, Pullman B *Quart. Rev. Biophys.*, **1987**, *20*, pp 113 - 172

13. Mitchell P R, Sigel H *Eur. J. Biochem.*, **1978**, *88*, pp 149 - 154

14. Reinhardt C G, Krugh T S *Biochemistry*, **1978**, *17*, N° 23, pp 4845 - 4854.

15. Veselkov A N, Djimant L N, Baranovsky S F *Chem. Phys.* (in Russian), **1989**, *8*, N° 9, pp 1282 - 1285

16. Nelder J, Mead R. *Comp. J.* **1965**, *7*, p. 308-311

17. Gusnin S J, Omeljaniov G A, Resnikov G V, Sirotkin V S. Minimisation in engineering calculations. *Moscow*, **1981**, pp.120.

18. Veselkov A N, Davies D B, Djimant L N, Parkes H G, *Molecular Biol. (in Russian), **1991**, 25 No. 5*, pp 1504-1516.

19. McGhee J D, von Hippel P H, *J. Mol. Biol.*, **1974**, *86*, pp 463 - 489

20. Delepierre M, Delbarre A, Destaintot B L, Igolen J *Biopolymers*, **1987**, *26*, pp 981 - 1000

21. Delepierre M, Hujinh Dinh T, Roques B P *Biopolymers*, **1989**, *28*, pp 2097 - 2113

22. Delepierre M, Hujinh Dinh T, Roques B P *Biopolymers*, **1989**, *28*, pp 2115 - 2142

23. Berman H M, Neidle S, Stodola R K *Proc. Natl. Acad. Sci. USA*, **1978**, *75*, N° 2, pp 828 - 832

24. Dickerson R E, *J. Biomol. Struct. Dyn.*, **1989**, *6* , 627-634.

25. Aggarwal A, Islam S A, Kuroda R, Neidle S *Biopolymers*, **1984**, *23*, pp 1025 - 1041

26. Neidle S, Pearl L N, Skelly J V *Biochem. J.*, **1987**, *243*, pp 1 - 13

Received 8/2/93
Accepted 11/11/93