

*Review Letter*

# Probing the three-dimensional structures of DNA and RNA oligonucleotides in solution by nuclear Overhauser enhancement measurements

G. Marius Clore and Angela M. Gronenborn

*Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG*

Received 9 November 1984

DNA and RNA oligonucleotides are ideally suited for high-resolution X-ray crystallographic and  $^1\text{H}$ -NMR studies. The solution structures of such oligonucleotides can potentially be solved using proton-proton nuclear Overhauser enhancement measurements to demonstrate the proximity of protons in space and to determine their separation, thereby enabling a comparison of the structure in the crystalline and solution states to be made. In this review we describe (i) the general strategy for the sequential resonance assignments of oligonucleotide  $^1\text{H}$ -NMR spectra, the essential prerequisite for further structural work, (ii) the approach to obtaining interproton distances from pre-steady state nuclear Overhauser enhancement measurements, and (iii) the use of interproton distances in structure determination. This is illustrated by several examples including double- and single-stranded DNA oligonucleotides as well as RNA stem and loop structures.

*Oligonucleotide    NMR    NOE    Interproton distance    Solution structure*

## 1. INTRODUCTION

It is only recently with the advent of fast and efficient methods of large-scale chemical synthesis of DNA and RNA fragments [1] that the crystal structures of a number of short oligonucleotides comprising examples of the A, B and Z families of DNA have been solved (see [2,3] for reviews). These studies have revealed the detailed arrangement of individual base pairs in the DNA helix at a resolution heretofore unavailable since fibre diffraction studies can only yield the averaged structure of a nucleic acid polymer [4]. However, although X-ray diffraction undoubtedly provides extensive structural detail, the crystal and fibre structures of oligonucleotide crystals and polynucleotide fibres are subject to crystal packing forces and local high ionic conditions [5]. These forces may well distort the structures and account for much of the local structure variations observed. It is therefore of considerable interest to develop methods for the determination of the three-dimensional solution structures of oligonucleotides

under physiological conditions where intermolecular forces are considerably weaker. Fortunately, short oligonucleotides are ideally suited for high resolution  $^1\text{H}$ -NMR studies. Consequently, their solution structures can potentially be solved, thereby enabling a comparison of the structures in the crystalline and solution states to be carried out. In this review, we will summarize the present state of the art with regard to achieving these aims. It should also be noted that  $^1\text{H}$ -NMR can yield considerable information on the dynamic properties of such oligonucleotides in solution, for example imino proton exchange rates. This area has been extensively reviewed recently by several groups [6-8] and will not be discussed here.

Essentially there are four NMR approaches which can be used to obtain structural information on an oligonucleotide in solution:

(i) The analysis of chemical shifts by calculations of through-space magnetic effects (ring current shift calculations). Although employed in a number of studies, the results are only qualitative

in nature and are based on available crystal structure data as input parameters [9,10].

(ii) The use of paramagnetic relaxation effects is well known, but its application is fraught with difficulties since it is based on a considerable number of assumptions, particularly when an external paramagnetic probe is used [11].

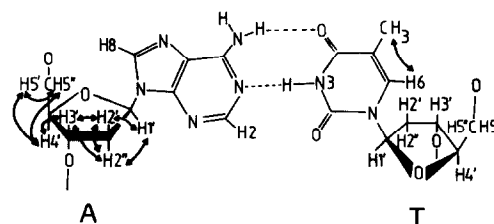
(iii) The analysis of three bond spin-spin coupling constants has been widely used to extract information concerning dihedral angles [12–15]. It suffers, however, from the fact that the relationship between three bond spin-spin coupling constants and dihedral angles is solely empirical in nature [11,16], and consequently the information obtained is essentially of a qualitative rather than a quantitative nature. A further point to consider if one wants to extract structural information from three bond proton-proton coupling constants is the fact that under conditions where small oligonucleotides of 6–12 base pairs are entirely double-stranded, namely at temperatures between 0 and 25°C, the appropriate coupling constants are difficult to resolve due to fairly large linewidths.

(iv) Potentially the most direct and powerful method of conformational analysis in solution is the use of the proton-proton nuclear Overhauser effect (NOE) which can demonstrate the proximity of two protons in space and can be used to determine their separation [17]. This approach has met with considerable success in the study of small proteins [18–25], ligand-protein [26–33], nucleic acid-nucleic acid [34,35] and drug-nucleic acid interactions [36–38], transfer ribonucleic acids [39–44], and oligonucleotides [45–59]. In most cases to date, NOE data have only been interpreted to yield qualitative structural information. With regard to nucleic acids this has proved particularly useful in examining the pattern of secondary and tertiary hydrogen bonding interactions in transfer ribonucleic acids [39–44], in distinguishing A, B and Z DNA [46,60,61], in examining the effects of base pair mismatching on nucleic acid conformation [62–63] and in monitoring intermolecular contact points in drug-oligonucleotide complexes [36–38]. However, a number of examples have recently appeared where quantitative structural information has been obtained [26–34, 47–53, 64–66].

## 2. SEQUENTIAL RESONANCE ASSIGNMENT

The full potential of NMR spectroscopy for structural studies can only be realised after identification of the individual resonance lines, and a general scheme for obtaining sequential assignments in protein  $^1\text{H}$ -NMR spectra has been described by Wüthrich and his collaborators [67]. In a similar manner, a sequential assignment method, limited to the imino and adenosine H2 proton resonances of adjacent base pairs in tRNA was developed by Redfield et al. [68]. Application of essentially the same principles to spectra of oligonucleotides can be used to achieve virtually complete resonance assignments [45–50, 54–59]. These assignments are a necessary prerequisite for the subsequent structure determination based on the quantitation of NOEs.

The assignment strategy involves basically two different NMR experiments. Firstly, it is helpful to identify the sugar spin system of a particular nucleotide either by decoupling or by two-dimensional  $J$  correlated spectroscopy (COSY). Thus the sugar resonances can be grouped into families of signals belonging to the same network of coupled spins via the intranucleotide pathway  $\text{H1}' \leftrightarrow \text{H2}' / \text{H2}'' \leftrightarrow \text{H3}' \leftrightarrow \text{H4}' \leftrightarrow \text{H5}' / \text{H5}''$  (fig.1). Since the



### COSY CONNECTIVITIES

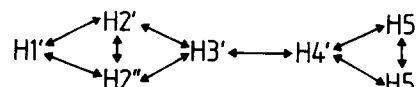
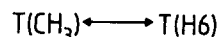
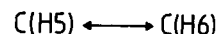


Fig. 1. Schematic representation of through-bond  $J$  connectivities in an AT base pair. All COSY connectivities in DNA are listed in the bottom part of the figure.

chemical shift dispersion of the H3' resonances is limited, the intranucleotide connectivity between the H3' and H4' resonances cannot be established unambiguously from the COSY spectrum alone, and it is therefore necessary to make use of NOE measurements to connect the H1' and H4' proton resonances of the same sugar residue on the basis of the close spatial proximity of the two atoms (see below). In addition  $J$  connectivities between the H5 and H6 proton resonances of cytosine residues and between the H6 and methyl proton resonances of thymine residues (via their four bond spin-spin coupling) are readily observed in the COSY spectrum.

As a second step, all protons that are separated by short distances ( $< 5 \text{ \AA}$ ) within the spatial structure can be connected by NOE measurements, either by one-dimensional pre-steady-state NOE measurements or by two-dimensional NOE spectroscopy (NOESY). Thus in double-stranded oligonucleotides, neighbouring bases can be identified as well as bases belonging to two different strands which are involved in base pairing. These NOE measurements provide the main body of information necessary for assignment, and in cases where not all couplings can be resolved, will lead to virtually complete assignments in their own right. Fig.2 summarizes a comprehensive NOE strategy for the assignment of all proton resonances in right-handed single- and double-stranded DNA helices. In the case of left-handed Z DNA, the intranucleotide distance relationships are the same as in the case of right-handed DNA (although the relative magnitudes of the sugar-base NOEs are significantly different for the purine residues which adopt a *syn* conformation as opposed to the *anti* conformation in B DNA). The internucleotide distance relationships, however, are entirely different in Z DNA and these are summarized in fig.3.

It is important to bear in mind that the initial assumption for the NOE-based assignment strategy of a particular helical structure and the subsequent refinement of the structure based on quantitative NOE data does not lead to pitfalls associated with a circular argument [45]. This is because of additional demands, constraints, and information extracted from  $J$  connectivities, the known nucleotide sequence, the nature of the terminal residues, and, most of all, the directionality of some of the internucleotide NOEs. Furthermore, in most cases a

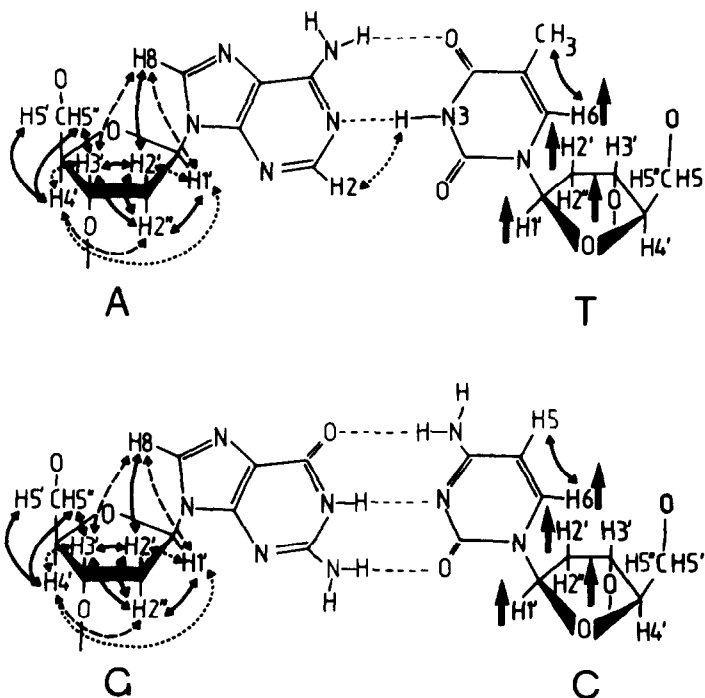
knowledge of the general conformational class of a particular oligonucleotide, A, B or Z, is readily obtained from a CD spectrum.

### 3. INTERPROTON DISTANCE DETERMINATION

In addition to providing assignments, NOE measurements can be used to determine interproton distances. For quantitation of the NOE we have used conventional one-dimensional NMR, although pure phase absorption NOESY experiments [69,70] with small random variations in the mixing time to eliminate zero quantum coherence transfer [71] may provide an alternative. The one-dimensional NOE experiment involves the saturation of the resonance of proton  $i$  and observing the intensity of the other proton resonances. For large molecules ( $M_r > 1000$ ) with long correlation times ( $\tau_c > 5 \times 10^{-10} \text{ s}$ ) such as oligonucleotides of 6 base pairs and longer, for which  $\omega\tau_c > 1$ , the NOEs observed are negative [72]. However, when  $\omega\tau_c \gg 1$ , the NOEs will no longer be selective in the steady state (i.e., following saturation of the resonance of proton  $i$  for  $t \rightarrow \infty$ ) owing to highly effective cross-relaxation between a large number of protons, a phenomenon known as spin diffusion [73], so that no structural information can be obtained. This problem can be completely circumvented by using only short times (typically  $\leq 0.5 \text{ s}$  for a molecule of  $M_r \sim 6000$ ) for either the selective saturation pulse in the one-dimensional experiment or for the mixing time in the two-dimensional experiment [74-77]. Under these conditions the pre-steady state NOE between two protons  $i$  and  $j$ ,  $N_{ij}$ , is given by

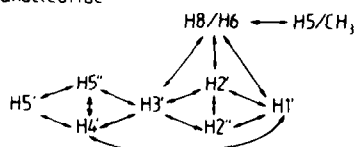
$$N_{ij}(t) \sim \sigma_{ij}t \quad (1)$$

providing  $\sigma_{ij} \geq \sigma_{ik}$  or  $\sigma_{ij} \geq \sigma_{jk}$  (where  $k$  is any other proton), as the initial build-up rate of the NOE is equal to the cross-relaxation rate  $\sigma_{ij}$  between the two protons  $i$  and  $j$  [74,75]. Distance information can then be obtained since  $\sigma_{ij}$  is inversely proportional to the sixth power of the distance,  $\langle r_{ij}^{-6} \rangle$ , between the two protons [72]. As a result of the  $\langle r_{ij}^{-6} \rangle$  dependence, the magnitude of the pre-steady-state NOE is very sensitive to interproton distance, decreasing rapidly as  $r_{ij}$  increases and becoming virtually undetectable for  $r_{ij} \geq 5 \text{ \AA}$ .

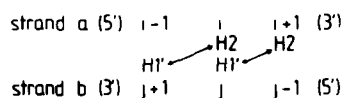


## A NOEs involving non-exchangeable protons

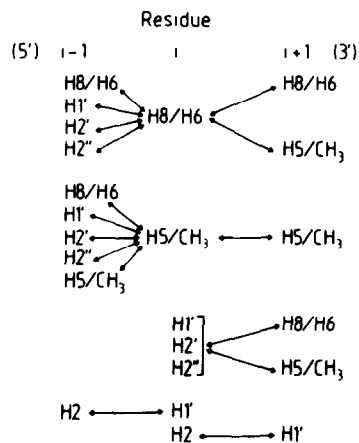
## 1 Intranucleotide



## 3 Internucleotide (interstrand)



## 2 Internucleotide (intrastrand)



## B NOEs involving exchangeable protons

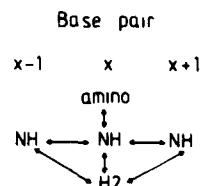


Fig.2. Schematic representation of through-space connectivities for right-handed B DNA. The intranucleotide interproton distances are represented as follows:  $\longleftrightarrow \leq 2.5 \text{ \AA}$ ;  $\cdots \cdots \leq 3.5 \text{ \AA}$ ;  $\cdots \cdots \leq 5 \text{ \AA}$ . Internucleotide distance relationships are shown on the right-hand side with large arrows ( $\uparrow$ ) next to those protons that are separated by  $\leq 5 \text{ \AA}$  from the H8, H6, H5 methyl protons of the base on their 3' side. All distance relationships are equally applicable to A DNA with the exception of the intranucleotide distance between the H2' and H8/H6 proton which is larger than  $3.5 \text{ \AA}$  and the intranucleotide distance between the H3' and H8/H6 protons which is less than  $3.5 \text{ \AA}$ . The bottom part of the figure lists all interproton NOE connectivities (distances  $\leq 5 \text{ \AA}$ ) which are applicable to both right-handed A and B DNA.

The ratio of two interproton distances may thus be obtained from the equation

$$(\langle r_{ij}^{-6} \rangle / \langle r_{kl}^{-6} \rangle)^{1/6} = (\sigma_{kl} / \sigma_{ij})^{1/6} \sim [N_{kl}(t) / N_{ij}(t)]^{1/6} \quad (2)$$

providing the correlation times for the two interproton distance vectors  $i-j$  and  $k-l$  are the same. If one of the distances is known, actual interproton distances can also be calculated. It should be noted that the approximation in eqn 2 remains valid up to values of  $t$  3-4-times longer than that in eqn 1 [78].

In the case of double-stranded oligodeoxyribonucleotides, there are three intranucleotide reference distances which are completely independent

of the DNA structure: namely,  $r_{H2'-H2''}$ ,  $r_{C(H6)-C(H5)}$ , and  $r_{T(H6)-T(CH_3)}$  which on the basis of standard bond lengths and angles have values of 1.8, 2.5 and  $2.7 \text{ \AA}$ , respectively. (Note that the latter distance is an average given by  $(\langle r_{ij}^{-6} \rangle)^{-1/6}$  calculated on the assumption of free rotation of the methyl group.) In addition, the distance between the T(H3) and A(H2) protons in a standard Watson-Crick base pair is  $2.9 \text{ \AA}$ ; although this distance will not be affected by propellor twisting it will be dependent on any deviation from idealized hydrogen bond length and geometry.

In order to make the appropriate choice of reference distance for the calculation of unknown interproton distances, one has to consider the expected ranges of the various interproton distances

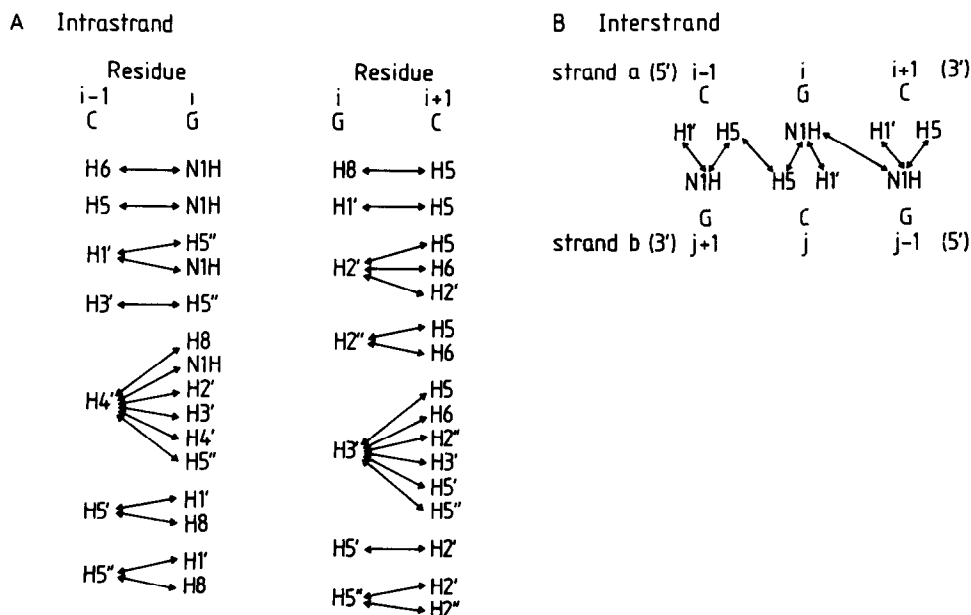


Fig.3. Internucleotide NOE connectivities for left-handed Z DNA [poly d(GC)]. All interproton distances  $\leq 5 \text{ \AA}$  are listed.

on the basis of stereochemical arguments and the possible motions of the different protons based on NMR relaxation studies and theoretical molecular dynamics calculations [79–88]. Taking these facts into account, it is reasonable to assume that [50]:

(i) the correlation times for the sugar-sugar, sugar-base (with the exception of the H1' sugar-base) and sugar-methyl interproton vectors are the same as those of the intranucleotide H2'-H2'' vector;

(ii) the correlation times for the base-base and the H1' sugar-base proton vectors are the same as those of the intranucleotide C(H5)-C(H6) vector; and

(iii) the correlation times for the base-methyl proton vectors are the same as those of the intranucleotide T(H5)-T(CH<sub>3</sub>) vector.

These assumptions are based on the fact that contributions from internal motion to the correlation times for case (i) will be dominated by motions within the sugar ring, for case (ii) by motions about the glycosidic bond and for case (iii) by rotation of the methyl group and motion about the glycosidic bond. A check on the validity of the above assumptions has been carried out by calculating various interproton distances whose idealized values are known using the above reference distances, resulting in calculated values all within 0.2 Å of the predicted ones [50]. Furthermore the calculations for case (iii) showed that the contribution from free rotation of the methyl group to the correlation time of the T(H5)-T(CH<sub>3</sub>) and base-methyl proton vectors is the same as that of the C(H5)-C(H6) vector within experimental error [50,87].

Using the procedure outlined above, a large set of intra- and internucleotide distances can be calculated from pre-steady-state NOE data, and given the interconvertibility of distances and cartesian coordinates [64–67, 88,89] these distances can in principle be used to determine the three-dimensional structure of an oligonucleotide. In the case of the four B type oligonucleotides on which this approach has been used, the overall root mean square difference between the NMR distances and those derived from fibre diffraction on B DNA is  $\leq 0.5$  Å [50,52,53].

#### 4. SOLUTION STRUCTURES

Because of the limited degrees of freedom available for a double-stranded oligonucleotide structure, one would expect that a reasonably large number of interproton distances would be sufficient to determine the three-dimensional solution structure with a high degree of confidence. In principle, these structures can be solved by manual model building, and indeed reasonably accurate values for the glycosidic  $\chi$  and C4'-C3' ( $\delta$ ) bond torsion angles can be obtained in this manner. However, because of potential cumulative errors inherent in such an approach, only qualitative information can be deduced for the other structural parameters, namely, backbone torsion angles, helical twist, helical rise and base tilt. This problem can potentially be overcome using a non-linear restrained least-squares optimization procedure in which all covalent bond lengths, fixed bond angles, van der Waals contacts, and hydrogen bond lengths and geometry are constrained within narrow limits, in order to refine an initial trial model on the basis of the interproton distance data determined from NOE measurements [90].

The minimum requirement to define both the glycosidic bond torsion angle ( $\chi$ ) and the sugar pucker conformation, defined in terms of the C4'-C3' bond torsion angle ( $\delta$ ), is two out of the three intranucleotide sugar-base distances,  $r_{\text{H1}'-\text{H8}/\text{H6}'}$ ,  $r_{\text{H2}'-\text{H8}/\text{H6}}$  and  $r_{\text{H3}'-\text{H8}/\text{H6}}$ . The *syn* and *anti* ranges for  $\chi$  are  $60 \pm 90^\circ$  and  $240 \pm 90^\circ$ , respectively. The distance  $r_{\text{H1}'-\text{H8}/\text{H6}}$  has a minimum value of 2.3–2.5 Å at  $\chi = 60^\circ$  (*syn*) and a maximum value of 3.7–3.9 Å at  $\chi = 240^\circ$  (*anti*). In addition, each value of  $r_{\text{H1}'-\text{H8}/\text{H6}}$  is compatible with two values of  $\chi$ :  $60^\circ < \chi_1 < 240^\circ$  and  $\chi_2 = (240^\circ - \chi_1) + 240^\circ$ . Given the restricted degrees of freedom imposed by the five-membered sugar ring, the distance  $r_{\text{H2}'-\text{H8}/\text{H6}}$  enables one to distinguish between  $\chi_1$  and  $\chi_2$  and to determine simultaneously the C4'-C3' bond torsion angle ( $\delta$ ). Similar arguments apply to other combinations of these three distances. Naturally, the more distances available, the better the determination of  $\chi$  and  $\delta$ , and in this respect the intranucleotide sugar-sugar interproton distance ratios  $r_{\text{H1}'-\text{H2}''}/r_{\text{H1}'-\text{H2}'}$  and  $r_{\text{H2}'-\text{H3}''}/r_{\text{H2}''-\text{H3}'}$  are quite helpful if available. In the case of B DNA where the conformation about the glycosidic bond is *anti* ( $\chi \sim -115 \pm 20^\circ$ ) and

the sugar pucker conformation resides in the  $O1'$ -*endo* to  $C2'$ -*endo* range, large intranucleotide NOEs will be observed between the H8/H6 and H2' protons and small intranucleotide NOEs between the H8/H6 and H1' protons and between the H8/H6 and H3' protons. In the case of A DNA, the glycosidic bond and sugar pucker conformations are low *anti* ( $\chi \sim -160 \pm 10^\circ$ ) and 3'-*endo*, respectively, so that small to medium-sized intranucleotide NOEs would be expected between the H8/H6 and H2' protons and large ones between the H8/H6 and H3' protons. In Z DNA, the pyrimidine residues have essentially the same glycosidic bond and sugar pucker conformations as in B DNA and will therefore exhibit the same pattern of intranucleotide NOEs; in contrast, the purine residues have a *syn* conformation about the glycosidic bond ( $\chi \sim 60-70^\circ$ ) and a 3'-*endo* sugar pucker so that large intranucleotide NOEs will be observed between the H8/H6 and H1' protons and small ones between the H8/H6 and H2'/H2'' protons.

The  $C4'-C5'$  bond torsion angle ( $\gamma$ ) can also be uniquely defined, providing two out of the three intranucleotide distances  $r_{H3'-H5'}$ ,  $r_{H4'-H5'}$  and  $r_{H4'-H5'}$  are known.

Once the glycosidic bond and sugar pucker conformations are known for each nucleotide, the inter-residue interproton distances enable one to define the position of each individual base pair with respect to its adjacent neighbours in terms of approximate values of the helical rise, helical twist and base tilt. The handedness of the helix can be deduced from the directional specificity of the internucleotide NOEs. In particular, internucleotide NOEs are observed between the methyl protons of T residues and the H8/H6 proton of the adjacent 5' but not 3' residue in a right-handed structure. Similarly, internucleotide NOEs are observed between the H1', H2' and H2'' sugar proton resonances of a given nucleotide and the H8/H6 and H5/methyl proton resonances of its adjacent 3' but not 5' neighbour in right-handed DNA. In B DNA, the internucleotide NOEs between the H2' and H2'' protons and the H8/H6 protons of the adjacent 3' residue are always much smaller than the intranucleotide NOE between the H2' and H8/H6 protons. In A DNA, however, the converse is true as the H2' proton is much closer to the H8/H6 proton of the 3' residue than to the H8/H6 proton of its own residue.

The first example of quantitative NOE measurements to derive structural information on oligonucleotides is that afforded by a study by Patel et al. [47] in which distances between adjacent adenine H2 protons of the  $d(ATTA) \cdot d(TAAT)$  fragment of the self-complementary dodecamer



On the basis of these distances they deduced the presence of base pair propellor twisting in agreement with the crystallographic findings of Dickerson and Drew [91].

To date we have examined the solution structures of several oligonucleotides on the basis of the approach outlined above principally by manual model building. These include the DNA hexamer  $5'-d(CGTACG)_2$  and DNA octamer  $5'-d(ACGCGCGT)_2$  which are both characterized by an alternating pyrimidine-purine sequence [48,49,90]. In the case of both oligonucleotides the overall solution structure is that of right-handed B DNA, namely, a right-handed helix with a helical rise of  $\sim 3.3$  Å, 10 base pairs per turn and the base pairs approximately perpendicular to the helix axis. In the case of  $5'-d(CGTACG)_2$ , subtle local structural variations associated with the pyrimidine and purine nucleotides are superimposed on the overall structure but the mononucleotide repeating unit is preserved (see fig.4). More interestingly, a restrained least-squares refinement of the structure on the basis of the interproton distance data [90] demonstrated that all the base pairs were propellor twisted and this was most marked for the GC base pairs at positions 2 and 5 in the sequence. These local structural variations are similar though less marked than in the crystal structure of the B DNA dodecamer  $5'-d(CGCGAATTCGCG)_2$  solved by Dickerson and Drew [91], and this difference may in part be due to the absence of strong intermolecular interactions, such as crystal packing forces, in solution. In contrast,  $5'-d(ACGCGCGT)_2$  has a clear alternating structure with a dinucleotide repeat, alternation occurring in the local helical twist and the glycosidic bond, sugar pucker and phosphodiester backbone conformations (see fig. 5). This alternating structure is in principle similar to that of wrinkled B DNA found in fibres of poly  $d(GC)$  [92] and the model of alternating B DNA proposed by Klug et al. [93]. The existence of both subtle and more dramatic local sequence-specific variations in the solution structures of DNA oligo-

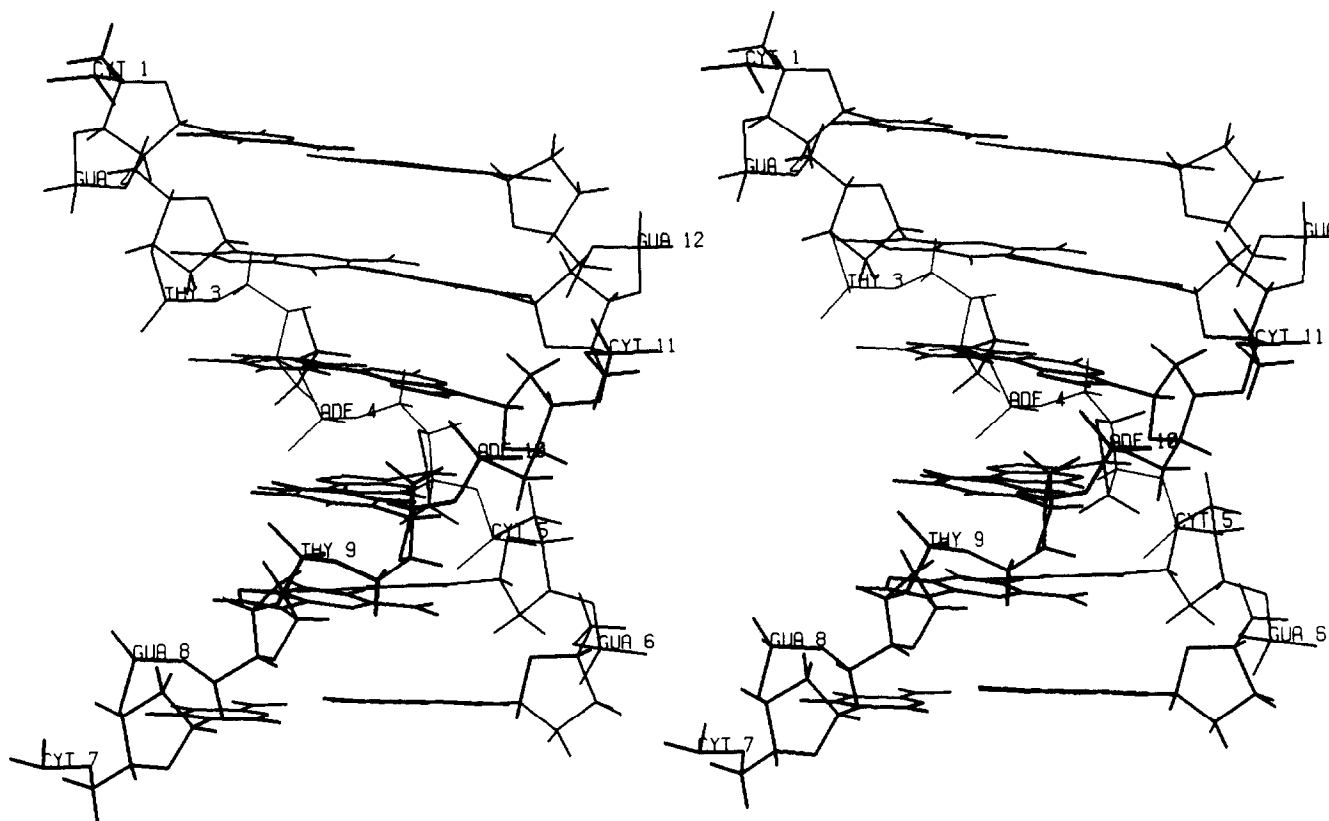


Fig.4. Stereo view of the refined solution structure of 5'-d(CGTACG)<sub>2</sub> [90]. This structure was generated by a restrained least-squares minimization [90] in which all covalent bond lengths, fixed bond angles, van der Waals contacts, and hydrogen bond lengths and geometry were constrained within narrow limits in order to refine an initial trial model on the basis of 200 interproton distances determined from NOE measurements [50,90].

nucleotides as exemplified by the hexamer and octamer, can clearly have a major influence on specific DNA-protein interactions. Thus, it is not difficult to visualize that the zig-zag distribution of the phosphorus atoms around a B DNA helix, as in the case of the octamer (fig.5), can present a specific multifaceted array of negative charges to a potentially interacting protein surface.

In addition to the hexamer and octamer discussed above, low-resolution structures of two non self-complementary DNA duplexes, namely the undecamer

[5'-d(AAGTGTGACAT) · 5'-d(ATGTCACACTT)] and the dodecamer [5'-d(CCAGAACAGTGG) · 5'-d(CCACTGTTCTGG)] comprising portions of the specific DNA target sites for the cyclic AMP receptor protein and the glucocorticoid receptor protein, respectively, have been determined [40,94].

Both these oligonucleotides were found to belong to the right-handed B DNA family. In the case of the undecamer, glycosidic ( $\chi$ ) and C4'-C3 ( $\delta$ ) bond torsion angles have also been obtained by model building on the basis of a large number of intranucleotide interproton distances [52]. It was found that whereas the sugar pucker exhibits little base to base variation with  $\delta$  lying in the range  $120 \pm 10^\circ$ , the glycosidic bond torsion angles of the pyrimidine and purine residues are significantly different with  $\chi_{\text{pyr}} \sim -120 \pm 10^\circ$  and  $\chi_{\text{pur}} \sim -90 \pm 10^\circ$  [52].

The strategy we have outlined is not just limited to double-stranded DNA structures. We have also examined the solution structure of the RNA pentadecamer 5'-r CAGAC<sub>m</sub>UG<sub>m</sub>AAYA/m<sup>5</sup>CUG comprising the anticodon loop and stem (residues 28-42) of yeast tRNA<sup>Phe</sup> (see fig.6), and carried



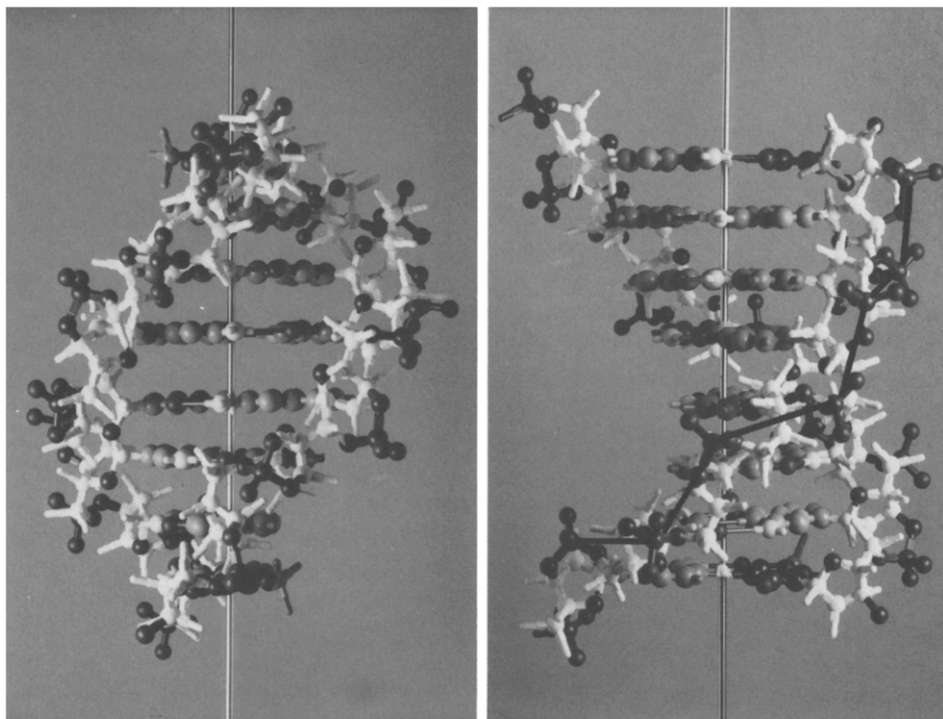


Fig.5. Two views of the solution structure of 5'-d(ACGCGCGT)<sub>2</sub> constructed from Nicholson models by manual model building on the basis of 79 interproton distances determined by NOE measurements [48,50]. The overall structure is that of B DNA. Superimposed on this structure is a dinucleotide repeating unit with alternation in the glycosidic bond, sugar pucker and backbone conformations. The zig-zag appearance of the phosphate backbone is highlighted.

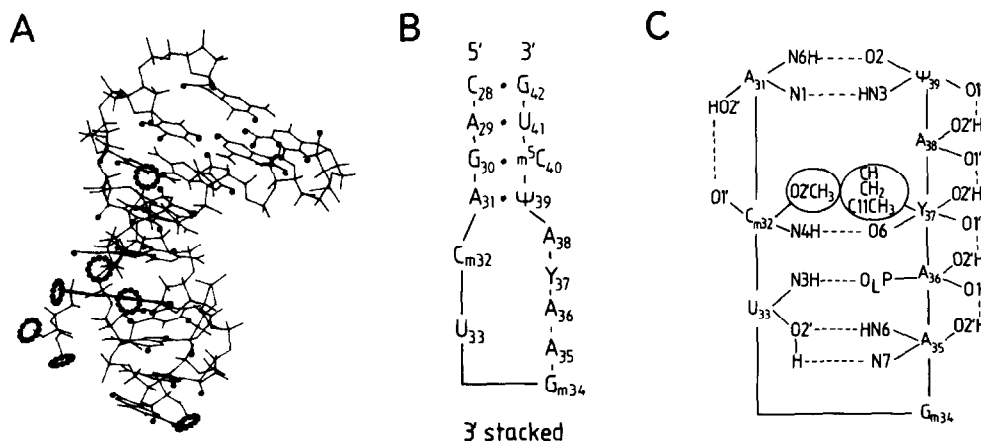


Fig.6. Diagrammatic representation of the solution structure of the pentadecamer 5'-r CAGAC<sub>m</sub>UG<sub>m</sub>AAAY<sub>m</sub>5CUG comprising the anticodon loop and stem of yeast tRNA<sup>Phe</sup> [51]. (A) Anticodon loop and stem structure as determined by X-ray diffraction of the monoclinic crystal form of yeast tRNA<sup>Phe</sup> [96]. Protons have been added using standard bond lengths and angles. The protons whose resonances have been assigned are highlighted. (B) Schematic representation of the 3' stacked hairpin loop structure as found in the crystal and in solution. (C) Stabilizing interactions for the 3' stacked loop conformation of the pentadecamer in solution deduced from model building on the basis of 75 interproton distances determined by NOE measurements [51]. Hydrogen bonds are represented by interrupted lines (---) and groups involved in hydrophobic interactions are encircled.

out a comparison of the solution structure with the crystallographic data of that portion of intact yeast tRNA<sup>Phe</sup> [50]. The pentadecamer adopts a hairpin loop structure in solution with the loop in a 3'-stacked conformation stabilized by both hydrogen bonding and hydrophobic interactions within the loop. The solution structure is both qualitatively and quantitatively remarkably similar to the crystal structure with an overall root mean square difference of only 1.2 Å between the interproton distances determined by NMR and X-ray crystallography.

Furthermore, single-stranded oligonucleotides which are inherently more flexible than their double-stranded counterparts are amenable to this approach, as has been shown for a single-stranded DNA undecamer [95]. The overall structure of this single-stranded DNA undecamer was found to be that of a right-handed B type helix with extensive base stacking. Within this overall structure there is quite a large degree of variability as exemplified by variations in glycosidic bond and sugar pucker conformation, most likely determined by base sequence.

In the examples given above, we have dealt only with isolated species. The same approach can also be used for interacting systems using transferred nuclear Overhauser enhancement (TRNOE) measurements [26,27]. These simply involve the extension of classical NOE measurements to exchanging systems, making use of chemical exchange to transfer magnetization concerning cross-relaxation between bound ligand protons from the bound to the free state. In this manner we have determined the solution structure of the ribotrinucleoside diphosphate rUpUpC, the codon for phenylalanine, bound to yeast tRNA<sup>Phe</sup> [34]. The glycosidic bond and ribose conformations are low *anti* and 3'-*endo*, respectively, typical of an A-RNA type structure. The main chain torsion angles are all within the range of those expected for A-RNA but small differences from those of conventional A-RNA 11 result in a special structure with a larger rotation per residue (40–45° compared to 32.7° in A-RNA 11) and almost perfect stacking of the bases. These two structural features, which are similar to those found in the anticodon triplet of the monoclinic form of yeast tRNA<sup>Phe</sup> [96], provide the underlying structural basis for the known greater stability of the codon-anticodon complex

relative to an equivalent double helical RNA trimer with a conventional A-RNA structure.

## 5. CONCLUDING REMARKS

Although NOE measurements in solution cannot compete as yet with single-crystal X-ray diffraction data, NMR solution data do provide important complementary information. This is particularly so when the different natures of the solution and crystal structures are borne in mind. NOE data are only capable of providing an interproton distance data set between protons separated by less than 5 Å, whereas high-resolution single X-ray diffraction is able to define the positions of all atoms in space with the exception of protons. Furthermore it might be interesting to investigate time-dependent fluctuations in the solution and crystal structures since they are reflected in the data obtained by both methods in a different manner. Whereas the crystal structure represents an average of linear superpositions of all the populations present in the crystal, distances obtained from NOE measurements in solution are  $(\langle r_{ij}^{-6} \rangle)^{-1/6}$  means. Thus, all fluctuations are equally weighted in the crystal structure, whereas the average distance between two protons in the solution structure is weighted in favour of fluctuations with shorter interproton distances.

In summary we would conclude that interproton distances obtained from NOE measurements will provide a powerful tool supplementing crystallographic studies, particularly in cases where crystal data are not available, in comparative studies of oligonucleotides with related sequences and in the study of transitions between different conformational states of nucleic acids.

## REFERENCES

- [1] Gassen, H.G. and Lang, A. (1982) Chemical and Enzymatic Synthesis of Gene Fragments, A Laboratory Manual, Verlag Chemie, Weinheim.
- [2] Dickerson, R.E., Drew, H.R., Conner, B.N., Kopka, M.L. and Pjura, P.E. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 13–24.
- [3] Wang, A.H.J., Fujii, S., Van Boom, J.H. and Rich, A. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 33–44.

- [4] Arnott, S., Chandrasekaran, R., Hall, I.H., Puigjaner, L.C., Walker, J.K. and Wang, M. (1983) Cold Spring Harbor Symp. Quant. Biol. 47, 53-66.
- [5] Rhodes, D. (1982) in: Topics in Nucleic Acid Structures Part 2, (Neidk, S. ed.) pp. 287-304, MacMillan, London.
- [6] Patel, D.J., Kozlowski, S.A., Ikuta, S., Itakura, K., Bhatt, R. and Hare, B.R. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 197-206.
- [7] Patel, D.J., Pardi, A. and Itakura, K. (1982) Science 216, 581-590.
- [8] Kearns, D.R. (1984) CRC Crit. Rev. Bioche. 15, 237-290.
- [9] Kan, L.S., Cheng, D.M., Jayaraman, K., Leutzinger, E.E., Miller, P.S. and Ts'o, P.O.P. (1982) Biochemistry 21, 6723-6732.
- [10] Bell, R.A., Alkema, D., Caddington, J.M., Hadder, P.A., Hughes, D.W. and Neilson, T. (1983) Nucleic Acids Res. 11, 1143-1149.
- [11] Jardetzky, O. and Roberts, G.C.K. (1981) NMR in Molecular Biology, Academic Press, New York.
- [12] Mellema, J.R., Haasnoot, C.A.G., Van Boom, J.H. and Altona, C. (1981) Biochim. Biophys. Acta 655, 256-264.
- [13] Tran-Dinh, S., Neuman, J.M., Huynh-Dinh, T., Gemissel, B., Igolen, J. and Simonot, G. (1982) Eur. J. Biochem. 124, 415-425.
- [14] Cheng, D.M., Kan, L.S., Leutzinger, E.E., Jayaraman, K., Miller, P.S. and Ts'o, P.O.P. (1982) Biochemistry 21, 621-630.
- [15] Altona, C. (1982) Red. Trav. Chim. Pays-Bas 101, 413-433.
- [16] Karplus, M. (1963) J. Am. Chem. Soc. 85, 2870-2871.
- [17] Noggle, J.H. and Schirmer, R.E. (1971) The Nuclear Overhauser Effect - Chemical Applications, Academic Press, New York.
- [18] Redfield, A.G. and Gupta, R.K. (1971) Cold Spring Harbor Symp. Quant. Biol. 36, 405-419.
- [19] Poulsen, F.M., Hosch, J.C. and Dobson, C.M. (1980) Biochemistry 19, 2596-2607.
- [20] Wagner, G., Kumar, A. and Wüthrich, K. (1981) Eur. J. Biochem. 114, 375-384.
- [21] Wagner, G. and Wüthrich, K. (1982) J. Mol. Biol. 159, 347-366.
- [22] Wagner, G. and Wüthrich, K. (1982) J. Mol. Biol. 160, 334-361.
- [23] Arseniev, S., Wider, G., Joubert, F.J. and Wüthrich, K. (1982) J. Mol. Biol. 159, 323-351.
- [24] Štrop, P., Wider, G. and Wüthrich, K. (1983) J. Mol. Biol. 166, 641-667.
- [25] Williamson, M.P., Marion, D. and Wüthrich, K. (1984) J. Mol. Biol. 173, 341-359.
- [26] Clore, G.M. and Gronenborn, A.M. (1982) J. Magn. Reson. 48, 402-417.
- [27] Clore, G.M. and Gronenborn, A.M. (1983) J. Magn. Reson. 53, 423-442.
- [28] Clore, G.M., Gronenborn, A.M., Mitchinson, C. and Green, N.M. (1982) Eur. J. Biochem. 128, 113-117.
- [29] Clore, G.M. and Gronenborn, A.M. (1982) FEBS Lett. 145, 197-202.
- [30] Gronenborn, A.M. and Clore, G.M. (1982) J. Mol. Biol. 157, 155-160.
- [31] Gronenborn, A.M. and Clore, G.M. (1982) Biochemistry 21, 4040-4048.
- [32] Gronenborn, A.M., Clore, G.M. and Jeffery, J. (1984) J. Mol. Biol. 172, 559-572.
- [33] Gronenborn, A.M., Clore, G.M., Brunori, M., Giardina, B., Falcioni, G. and Perutz, M.F. (1984) J. Mol. Biol. 178, 731-742.
- [34] Clore, G.M., Gronenborn, A.M. and McLaughlin, L.W. (1984) J. Mol. Biol. 174, 163-173.
- [35] Gronenborn, A.M., Clore, G.M., McLaughlin, L.W., Graeser, E., Lorber, B. and Giegé, R. (1984) Eur. J. Biochem., in press.
- [36] Patel, D.J. (1982) Proc. Natl. Acad. Sci. USA 79, 6424-6428.
- [37] Reid, D.G., Salisbury, S.A. and Williams, D.H. (1983) Biochemistry 22, 1377-1385.
- [38] Brown, S.C., Mullis, K., Levenson, C. and Shafer, R.H. (1984) Biochemistry 23, 403-408.
- [39] Johnston, P.D. and Redfield, A.G. (1978) Nucleic Acids. Res. 5, 3913-3927.
- [40] Roy, S. and Redfield, A.G. (1983) Biochemistry 22, 1386-1390.
- [41] Hare, D.R. and Reid, B.R. (1982) Biochemistry 21, 1835-1842.
- [42] Heerschap, A., Haasnoot, C.A.G. and Hilbers, C.W. (1982) Nucleic Acids Res. 10, 6981-7000.
- [43] Heerschap, A., Haasnoot, C.A.G. and Hilbers, C.W. (1983) Nucleic Acids Res. 11, 4483-4500.
- [44] Heerschap, A., Haasnoot, C.A.G. and Hilbers, C.W. (1983) Nucleic Acids Res. 11, 4501-4520.
- [45] Reid, D.G., Salisbury, S.A., Bellard, S., Shakked, Z. and Williams, D.H. (1983) Biochemistry 22, 2012-2025.
- [46] Reid, D.G., Salisbury, S.A., Brown, T., Williams, D.H., Vasseur, J.J., Rayner, B. and Imbach, J.L. (1983) Eur. J. Biochem. 135, 307-314.
- [47] Patel, D.J., Kozlowski, S.A. and Bhatt, R. (1983) Proc. Natl. Acad. Sci. USA 80, 3908-3912.
- [48] Clore, G.M. and Gronenborn, A.M. (1983) EMBO J. 2, 2109-2115.
- [49] Clore, G.M. and Gronenborn, A.M. (1984) Eur. J. Biochem. 141, 119-129.
- [50] Gronenborn, A.M., Clore, G.M. and Kimber, B.J. (1984) Biochem. J. 221, 723-736.

- [51] Clore, G.M., Gronenborn, A.M., Piper, E.A., McLaughlin, L.W., Graesser, E. and Van Boom, J.H. (1984) *Biochem. J.* 221, 735-750.
- [52] Gronenborn, A.M., Clore, G.M., Jones, M.B. and Jiricny, J. (1984) *FEBS Lett.* 165, 216-222.
- [53] Clore, G.M. and Gronenborn, A.M. (1984) *FEBS Lett.* 175, 117-123.
- [54] Sheck, R.M., Russo, N., Boelens, R., Kaptein, R. and Van Boom, J.H. (1983) *J. Am. Chem. Soc.* 105, 2914-2916.
- [55] Scheek, R.M., Boelens, R., Russo, N., Van Boom, J.H. and Kaptein, R. (1984) *Biochemistry* 23, 1371-1376.
- [56] Hare, D.R., Wemmer, D.E., Chou, S.H., Drobny, G. and Reid, B.R. (1983) *J. Mol. Biol.* 171, 319-336.
- [57] Weiss, M.A., Patel, D.J., Sauer, R.T. and Karplus, M. (1984) *Proc. Natl. Acad. Sci. USA* 81, 130-134.
- [58] Feigon, J., Leupin, W., Denny, W.A. and Kearns, D.R. (1983) *Biochemistry* 22, 5943-5951.
- [59] Frechet, D., Cheng, D.M., Kan, L.S. and Ts'o, P.O.P. (1983) *Biochemistry* 22, 5194-5200.
- [60] Patel, D.J., Kozlowski, S.A., Nordheim, A. and Rich, A. (1982) *Proc. Natl. Acad. Sci. USA* 79, 1413-1417.
- [61] Assa-Munt, N. and Kearns, D.R. (1984) *Biochemistry* 23, 791-796.
- [62] Patel, D.J., Kozlowski, S.A., Ikuta, S. and Itakura, K. (1984) *Biochemistry* 23, 3207-3217.
- [63] Patel, D.J., Kozlowski, S.A., Ikuta, S. and Itakura, K. (1984) *Biochemistry* 23, 3218-3226.
- [64] Braum, W., Bösch, C., Brown, L.R., Go, N. and Wüthrich, K. (1981) *Biochim. Biophys. Acta* 667, 377-396.
- [65] Braun, W., Wider, G., Lee, K.H. and Wüthrich, K. (1983) *J. Mol. Biol.* 169, 921-948.
- [66] Arseniev, S.A., Kondakov, V.I., Maiorov, V.N. and Bystrov, V.F. (1984) *FEBS Lett.* 165, 57-62.
- [67] Wüthrich, K., Wider, G., Wagner, G. and Braun, W. (1982) *J. Mol. Biol.* 155, 311-319.
- [68] Redfield, A.G., Roy, S., Sanchez, U., Tropp, J. and Figueroa, N. (1981) in: *Biomolecular Stereodynamics*, vol. I (Sarma, R.H. ed.) pp. 195-208, Adenine Press, New York.
- [69] States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982) *J. Magn. Reson.* 48, 286-292.
- [70] Marion, D. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* 113, 967-974.
- [71] Macura, S., Huang, Y., Sutter, D. and Ernst, R.R. (1981) *J. Magn. Reson.* 43, 259-281.
- [72] Solomon, I. (1955) *Phys. Rev.* 90, 559-565.
- [73] Kalk, A. and Berendsen, H.J.C. (1976) *J. Magn. Reson.* 24, 343-366.
- [74] Wagner, G. and Wüthrich, K. (1979) *J. Magn. Reson.* 33, 675-680.
- [75] Dobson, C.M., Olejniczak, E.T., Poulsen, F.M. and Ratcliffe, J. (1982) *J. Magn. Reson.* 48, 87-110.
- [76] Kumar, A., Wagner, G., Ernst, R.R. and Wüthrich, K. (1981) *J. Am. Chem. Soc.* 103, 3654-3658.
- [77] Macura, S. and Ernst, R.R. (1980) *Mol. Phys.* 41, 95-117.
- [78] Clore, G.M. and Gronenborn, A.M. (1984) *J. Magn. Reson.*, in press.
- [79] Hogan, M.E. and Jardetzky, O. (1979) *Proc. Natl. Acad. Sci. USA* 76, 6341-6345.
- [80] Bolton, P.H. and James, T.L. (1979) *J. Chem. Phys.* 83, 3359-3367.
- [81] Shindo, H. (1980) *Biopolymers* 19, 509-522.
- [82] Lipari, G. and Szabo, A. (1981) *Biochemistry* 20, 6520-6526.
- [83] Kollman, P., Keepers, J.W. and Weiner, P. (1982) *Biopolymers* 21, 2345-2376.
- [84] Levitt, M. (1983) *Cold Spring Harbor Symp. Quant. Biol.* 47, 251-262.
- [85] Tidor, B., Irikura, K., Brooks, B.R. and Karplus, M. (1983) *J. Biomol. Struct. Dynamics* 1, 231-252.
- [86] Holbrook, S.R. and Kim, S.H. (1984) *J. Mol. Biol.* 173, 361.
- [87] Clore, G.M. and Gronenborn, A.M. (1984) *FEBS Lett.* 172, 219-225.
- [88] Crippen, G.M. and Havel, T.F. (1978) *Acta Crystallogr. A* 34, 282-284.
- [89] Cohen, F.E. and Sternberg, M.J.E. (1980) *J. Mol. Biol.* 138, 321-333.
- [90] Clore, G.M., Gronenborn, A.M., Moss, D.S. and Tickle, I.J. (1985) *J. Mol. Biol.*, in press.
- [91] Dickerson, R.E. and Drew, H.R. (1981) *J. Mol. Biol.* 149, 761-786.
- [92] Arnott, S., Chandrasekaran, R., Puigjaner, L.C., Walker, J.K., Hall, I.H., Birdsall, D.L. and Ratcliff, R.L. (1983) *Nucleic Acids Res.* 11, 1457-1474.
- [93] Klug, A., Jack, A., Viswamitra, M.A., Kennard, O., Shakked, Z. and Steitz, T.A. (1979) *J. Mol. Biol.* 131, 669-680.
- [94] Clore, G.M., Lauble, H., Frenkiel, T.A. and Gronenborn, A.M. (1984) *Eur. J. Biochem.*, in press.
- [95] Clore, G.M. and Gronenborn, A.M. (1984) *Eur. Biophys. J.* 11, 95-102.
- [96] Jack, A., Lardner, J.E. and Klug, A. (1976) *J. Mol. Biol.* 108, 619-649.