Biochemistry

 Volume 26, Number 21 October 20, 1987

Perspectives in Biochemistry

Solution Conformations of N-Linked Oligosaccharides†

S. W. Homans,* R. A. Dwek, and T. W. Rademacher

Oxford Oligosaccharide Group, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, England

Received June 1, 1987; Revised Manuscript Received July 15, 1987

An integral feature of many cell surface and secreted proteins is the presence of one or more oligosaccharide units covalently linked to asparagine. Despite intensive investigation, the function of these oligosaccharides is in general unknown, although they have been implicated in a variety of biological systems. Of particular interest is their proposed role as recognition signals in such processes as intracellular migration and secretion of glycoproteins (Sharon & Lis, 1982), biological activity of glycohormones (Calvo & Ryan, 1985), clearance of glycoproteins from circulation by hepatocytes (Morell et al., 1968; Prieels et al., 1978; Hudgin et al., 1974), and the metastasis of tumor cells (Dennis & Laferté, 1985). As a first step toward an understanding of these and other postulated functions, attention has been focused upon oligosaccharide three-dimensional structure. High-resolution ^{1}H NMR[1] has been the primary tool for these investigations (Homans et al., 1982; Bock et al., 1982; Brisson & Carver, 1983a,b; Paulsen et al., 1984, 1986).

The purpose of this review is to show how the current understanding of oligosaccharide solution conformations and dynamics has been furthered by recent developments in NMR methodology together with the use of energy calculations. First, a general strategy for the assignment of ^{1}H NMR spectra of oligosaccharides is described. The use of these assignments for the interpretation of nuclear Overhauser effect (NOE) experiments is then discussed. The latter generate distance constraints, which, in conjunction with conformational energy calculations and molecular dynamics simulations, allow the structural and dynamic properties of a given oligosaccharide to be determined. Finally, we explore the extent to which the proposed recognition function of oligosaccharides can be rationalized in terms of these properties.

Assignment Methods

A representative oligosaccharide primary structure is shown in Figure 1. In spite of their relatively small size, proton NMR spectra of oligosaccharides are surprisingly complex. The ^{1}H NMR spectrum of the pentasaccharide core common to most Asn-linked oligosaccharides is shown in Figure 2. Only the anomeric and mannosyl H2 protons are well resolved at low field due to the electron-withdrawing properties of the ring oxygen. The remaining proton resonances, with few exceptions (e.g., *N*-acetamido methyl and fucose methyl groups), resonate in the unresolved envelope between ~3.5 and 4.0 ppm and are often strongly coupled. Since an essentially complete set of unambiguous proton resonance assignments is required for conformational analysis, an effective assignment strategy is required.

The resonance positions of the resolved anomeric protons of a given oligosaccharide are sensitive to the spatial disposition of the constituent monosaccharides and hence also to the primary sequence. These resonances may thus be assigned with reference to suitable model compounds. This approach has been used by Vliegenthart et al., mainly in the assignment of anomeric and mannosyl H2 protons. These workers have also demonstrated the use of these assignments for primary sequence determination in oligosaccharides [see Vliegenthart et al. (1983) for a review].

An important advance for the determination of resonance assignments in oligosaccharides lies in the application of 2D NMR techniques that have been so successful in the assignment of the ^{1}H NMR spectra of proteins (Wüthrich et al., 1982). The availability of unambiguous anomeric and mannosyl H2 proton assignments suggests an obvious assignment

† This work is a contribution from the Oxford Oligosaccharide Group, which is supported by Monsanto. The authors are members of the Oxford Enzyme Group.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; COSY, ^{1}H–^{1}H correlated spectroscopy; TQCOSY, triple quantum filtered COSY; NOESY, two-dimensional (2D) ^{1}H–^{1}H NOE spectroscopy; RECSY, multistep relayed correlation spectroscopy; MNDO, modified neglect of diatomic differential overlap.

0006-2960/87/0426-6571$01.50/0

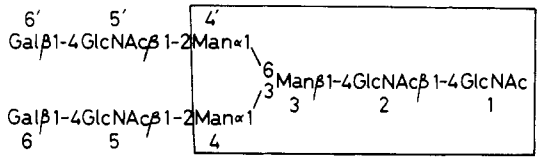

FIGURE 1: Representative oligosaccharide primary structure, the biantennary N-linked oligosaccharide from human serum transferrin. *Note on nomenclature*: the pentasaccharide core common to most N-linked oligosaccharides is shown boxed. Substitution on the α-mannosyl residues of the core gives rise to the α1–6 and α1–3 antennae.

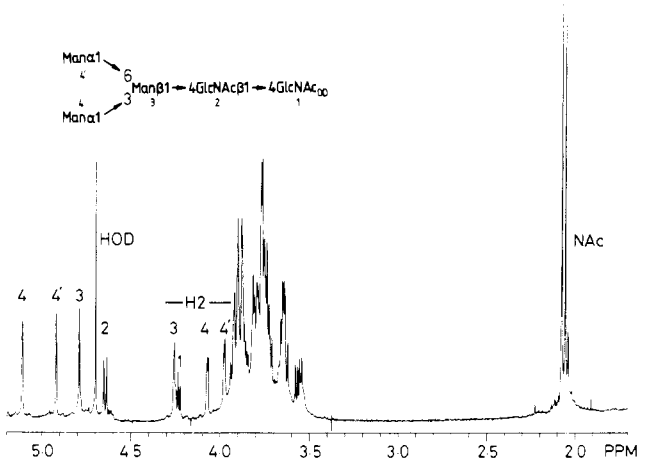

FIGURE 2: One-dimensional 500-MHz ^{1}H NMR spectrum of the common pentasaccharide core (reduced form). The H1 and mannosyl H2 protons resonate at low field due to the electron-withdrawing properties of the ring oxygen. The remaining protons (with the exception of *N*-acetamido methyl protons) resonate in the region ~3.5–4.0 ppm.

strategy. By use of ^{1}H–^{1}H correlated spectroscopy (COSY; Bax & Freeman, 1981), it should be possible in theory to trace the through-bond (scalar) coupling networks around each monosaccharide ring, using the anomeric proton assignments as a basis (Figure 3a). Since there is negligible spin–spin (*J*) coupling between H1 and H5, there is no ambiguity in the direction in which the connectivities should be traced around the ring. As an example, a section of the ^{1}H–^{1}H COSY spectrum of the pentasaccharide core is shown in Figure 4a. The region between 3.3 and 4.3 ppm has been chosen to emphasize the practical difficulties with this approach. For example, while the correlation between mannosyl H1's and H2's (not shown) and mannosyl H2's and H3's are easily detected, the degree of resonance (cross-peak) overlap that is still apparent in the 2D spectrum prevents unambiguous assignments of mannosyl H4's and thus also the remaining mannosyl protons. Occasionally, the situation is simplified if a resolved coupling exists between next-neighbor protons. A fortuitous example of this is seen in the cross-peak connecting Man-4 H2 with Man-4 H4. As an aside, note that strong coupling can be detected in Man-3 between H3 and H4. This is apparent (and diagnostic) from the "skewed" intensity of the cross-peak correlating Man-3 H2 with Man-3 H3/H4.

The resonance overlap problem is no less significant for residues such as β-D-GlcNAc and β-D-galactose in larger oligosaccharide structures. Consequently, it is necessary to supplement the COSY technique with additional methods in order to obtain complete resonance assignments (Homans et al., 1986). In choosing these methods, the overall aim must be either to reduce the number of cross-peaks present such that resonance overlap does not occur or to obtain through-bond correlations from the H1 protons by using a fundamentally

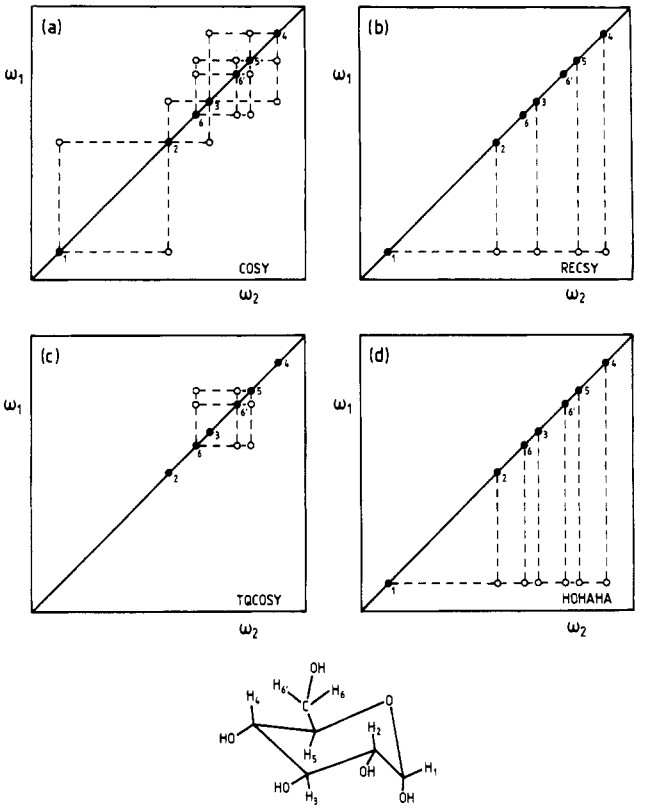

FIGURE 3: Diagrammatic illustration of two-dimensional NMR spectra resulting from the application of the various techniques described in the text. The COSY technique (a) correlates all protons between which a resolved scalar (*J*) coupling exists in each monosaccharide residue. This results in severe cross-peak overlap. In contrast, the RECSY technique (b) allows correlations to be made in a well-resolved region of the spectrum. However, uncertainties in the values of J_{56} and $J_{56'}$ (see text) do not allow correlations from H1 to H6 and H1 to H6′ to be made with confidence. The TQCOSY technique (c) is designed to overcome this difficulty, since the conventional COSY spectrum (a) is purged by the filter of all cross-peaks except those correlating H5 with H6, H5 with H6′, and H6 with H6′. The HOHAHA technique (d) is a more recent method by which complete monosaccharide residue subspectra can be extracted from the spectrum of the oligosaccharide. This arises from the complete correlation of H1 and the remaining ring protons independent of spin coupling topography.

different technique. With regard to the latter possibility, the application of multistep relayed correlation spectroscopy (RECSY) is of value (Eich et al., 1982; Homans et al., 1984). In this technique, rather than obtaining through-bond connectivities in a pairwise manner (i.e., H1–H2, H2–H3, H3–H4, etc.), cross-peaks appear correlating H1–H2, H1–H3, H1–H4, and H1–H5, despite vanishing couplings between next-neighbor protons. The technique works by relaying magnetization derived from H1 around the ring to H5 (Homans et al., 1984). The basic differences between COSY and RECSY are illustrated in Figure 3. The RECSY technique is of particular value in the assignment of NMR spectra of oligosaccharides that possess a degree of symmetry and/or repeating structure.

One difficulty associated with the RECSY experiment is that for a given experimental regime magnetization transfer is highly sensitive to the spin-coupling topography or "*J*-signature" of each monosaccharide residue. The consequences of this are twofold. First, if strong coupling is present within the spin system, then efficient transfer is unlikely to take place past the strongly coupled protons. The second difficulty derives from attempts to extend the RECSY sequence to encompass correlations of H1 with H6 and H1 with H6′. Since various

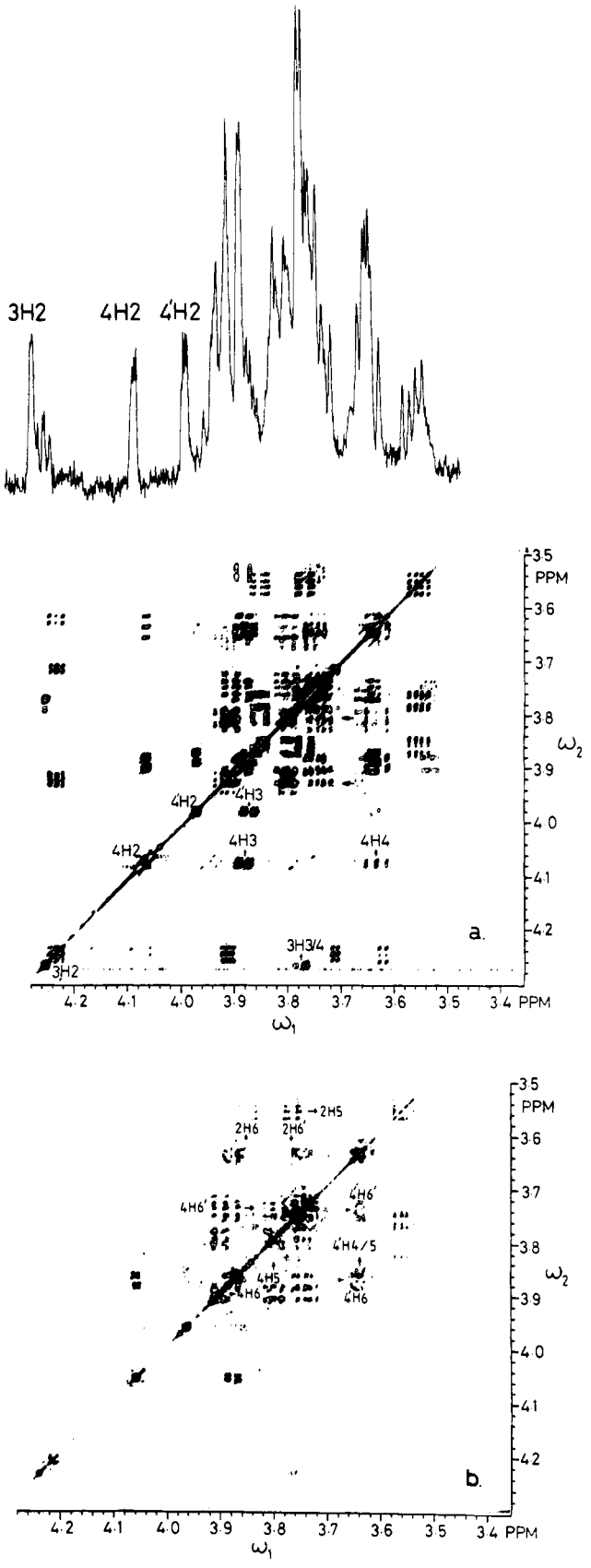

FIGURE 4: (a) ^1H–^1H COSY spectrum of the pentasaccharide core (Figure 1). The region 3.4–4.3 ppm has been chosen to emphasize assignment difficulties due to cross-peak overlap. (b) TQCOSY spectrum of the pentasaccharide core. This spectrum contains only cross-peaks correlating H5, H6, and H6′ protons. An exception is the cross-peak correlating Man-4 H2 with Man-4 H3 (see Figure 1 for nomenclature). This passes the filter due to resolved coupling between Man-4 H2 and Man-4 H4.

rotamer distributions are possible about the C5–C6 bond of a monosaccharide, which may be different in the oligosaccharide (Homans et al., 1986), the precise values of J_{56} and J_{56}' are not known with certainty. Unfortunately, these are required to compute the correct experimental conditions. For this reason a third 2D NMR technique for resonance assignments in oligosaccharides has been utilized. This method, known as triple quantum filtered ^1H–^1H correlated spectroscopy (TQCOSY) (Piantini et al., 1982), belongs to a class of techniques that achieve spectral simplification via spin filtering. Remembering that the H1–H5 proton resonance assignments are already available and that the conventional COSY spectrum contains *all* of the through-bond connectivities with resolved J couplings, it would be convenient to purge the COSY spectrum of all cross-peaks except those correlating H5–H6, H5–H6′, and H6–H6′, thus greatly reducing the cross-peak overlap problem. The purpose of TQCOSY is to achieve just this (Homans et al., 1986). The basic principles of TQCOSY are shown in Figure 3c [see Homans et al. (1986) for a more detailed description]. The key concept is that the H5, H6, and H6′ protons have mutual, resolved couplings, whereas the remaining ring protons do not. While the only requirement for the generation of triple quantum coherence in a three-spin system is two resolved couplings, a third coupling is required under which the spin system can evolve in order to generate a cross-peak in the TQCOSY spectrum. The final result is quite satisfactory, as shown in Figure 4b for the pentasaccharide core structure.

To complete the assignment strategy, it is necessary to examine those cases where strong coupling is present in the spin system(s) of one or more constituent monosaccharides. For these situations the application of ^1H–^1H homonuclear Hartmann–Hahn spectroscopy (HOHAHA) is of value (Davis & Bax, 1985; Homans et al., 1987a). There are important theoretical differences between HOHAHA and the three techniques described above. The most important of these differences lies in the mechanism by which coherence transfer occurs between the spins, which is in turn responsible for the formation of the cross-peaks that are necessary for the shift correlation. In HOHAHA, coherence transfer essentially involves the generation of strong coupling between all spins in a given coupling network (Boyd & Bazzo, 1987). The result is similar in RECSY in that correlations between H1 and H2, H1 and H3, H1 and H4, etc. are generated (Figure 3), but the coherence-transfer process is far less dependent upon the J-signature in HOHAHA than in RECSY, and it is possible to generate a complete monosaccharide subspectrum (i.e., including correlations between H1 and H6 and H1 and H6′) from cross sections through HOHAHA spectra even in the presence of inherent strong coupling (Homans et al., 1987a).

In summary, complete resonance assignments can be obtained for oligosaccharides with a combination of four techniques (COSY, RECSY, TQCOSY, HOHAHA). The HOHAHA method, in particular, shows great promise in that resonance assignments for an oligosaccharide are potentially available in a *single experiment*. However, techniques such as RECSY are still of value when it is necessary to maintain control over the coherence-transfer pathway. Such control is necessary, for example, in the assignment of oligosaccharides when attached to large peptides (~20 amino acids), in order to suppress cross-peaks derived from amino acid residues.

Determination of Primary Sequence

The NMR "fingerprint" method of oligosaccharide primary sequence determination has been proposed earlier. From the

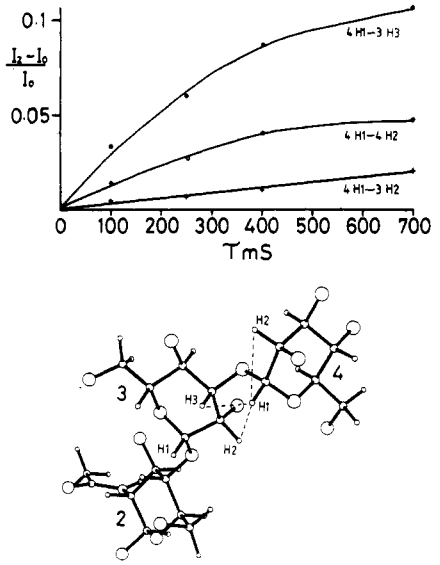

FIGURE 5: Typical time development of cross-peak intensities for the Manα1–3Manβ linkage of the pentasaccharide core. The initial buildup rates are proportional to the cross-relaxation rates between spin pairs, which are in turn proportional to the inverse sixth power of the internuclear distance. Distance information across glycosidic linkages may thus be computed by comparing the buildup rates of interresidue vs intraresidue NOEs, since the latter correspond to fixed internuclear distances.

careful 1D studies of Vliegenthart et al., a list of NMR data has been accumulated for a series of N- and O-linked oligosaccharides (Vliegenthart et al., 1983). The value of this approach as a rapid, simple, noninvasive sequence method should not be overlooked. However, a fundamental difficulty arises when a *completely* novel structure is under investigation. Under these circumstances an ab initio approach involving NOE data (see below) can be of value. However, often NOE data are not forthcoming because of low sample concentrations together with unmeasurably small NOE values due to the correlation time dependence of this parameter. In such cases, a third approach may be of value. It involves a comparison of the assignments of an oligosaccharide (from one or more of the techniques described above) with those that are readily available for the constituent monosaccharides in their free form. The chemical shift differences can then be interpreted in terms of linkage position, thus yielding information on three-dimensional structure. This approach has been used recently in the determination of the primary sequence of a novel class of glycan isolated from trypanosomes (Ferguson et al., 1987).

Determination of Three-Dimensional Structure

The availability of a complete set of assignments makes the determination of three-dimensional structure straightforward. The method consists essentially of the application of ^1H–^1H nuclear Overhauser effect spectroscopy (NOESY) (Muller & Ernst, 1980; States et al., 1982) to generate short-range through-space (<4 Å) connectivities within the molecule. The time development of the NOE, together with the use of intraresidue NOEs for calibration purposes, allows interresidue NOEs to be quantified (Clore & Gronenborn, 1985). Since each monosaccharide residue has a fixed ring geometry (which may be confirmed by measurement of vicinal coupling constants), the quantitation of internuclear distances across glycosidic linkages by following the time development of cross-peak intensities (Figure 5) allows the complete solution conformation of an oligosaccharide to be determined (Homans et al., 1987b). As an aside we note that a qualitative interpretation of the NOEs together with a complete set of resonance assignments is sufficient to generate ab initio the primary sequence of an unknown oligosaccharide without reference to model compounds, since NOE connectivities can be traced across glycosidic linkages. However, since NOE depends on conformation, care must be taken to ensure that the observed NOE corresponds to that from the anomeric proton to the aglycon (i.e., the proton on the carbon to which the glycosidic linkage is attached), rather than a proton adjacent to the latter.

Despite the apparent simplicity of NOESY, an important caveat remains. It cannot be assumed that on the basis of NOE data between *two* protons across a given glycosidic linkage that the computed solution conformation is unique, since the magnitude of the NOE is proportional to the inverse sixth power of the internuclear distance. While this renders NOE measurements particularly sensitive to small conformational changes, it also has the effect of artificially weighting conformers with close approach of two given nuclei in situations where more than one conformer exists. In other words, the NOE-derived structure may represent a conformation with close approach of two protons that is only fractionally coupled in favor of an alternative conformer where these protons are distant. For this reason it is important to acquire as many through-space connectivities as possible across glycosidic linkages.

Energy Calculations

The possibility of conformational averaging about a given glycosidic linkage can be assessed by computing a potential surface with respect to the torsional angles ϕ and ψ. In the absence of conformational averaging, the conformation corresponding to the computed energy minimum should agree closely with the NMR-derived conformation (Homans et al., 1987b,c).

The choice of method for the evaluation of conformational energies to some extent depends upon the size of the structure (number of degrees of freedom) to be investigated. In general, empirical methods such as hard-sphere exoanomeric effect (Bock et al., 1982; Brisson & Carver 1983a,b; Paulsen et al., 1984, 1986) and molecular mechanics (Lindon et al., 1984) have been the methods of choice for conformational analysis in oligosaccharides. The advantage of these techniques is that energy minimization of a complete oligosaccharide can be achieved at modest computational cost. In contrast, molecular orbital (MNDO) methods (Dewar & Thiel, 1977) are preferable to purely empirical approaches, since the latter can give a misleading result under certain circumstances (Homans et al., 1986). However, the use of quantum mechanical methods at present restricts energy calculations to disaccharide and trisaccharide fragments of oligosaccharides, in view of their computational complexity. While it is clear that this approach does not account for long-range interactions in oligosaccharides, such interactions, if they exist, can only *reduce* the number of available conformers. In this respect the value of MNDO methods is to determine whether more than one energy minimum exists on the potential surface and to gain some measure of the depth(s) (Homans et al., 1986, 1987b,c). Unfortunately, many investigations on oligosaccharide conformations appear to emphasize the results of the energy minimization by interpreting the NMR data in terms of it, rather than vice versa. At best, all energy minimization methods are an approximation to the true potential surface, particularly since the majority are performed in vacuo. The

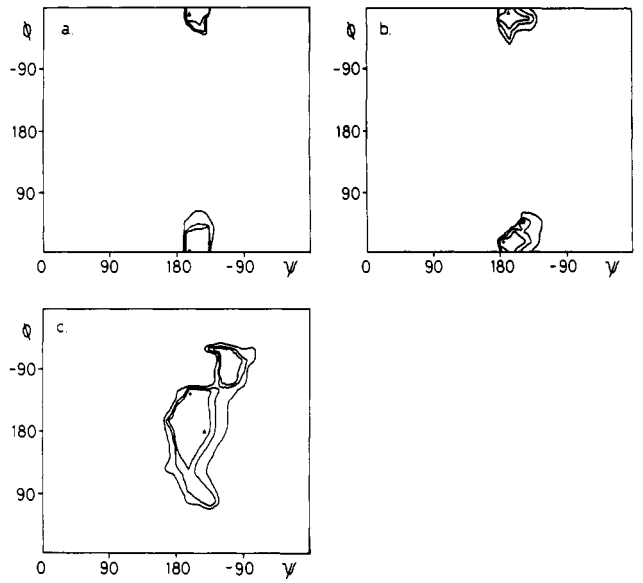

FIGURE 6: Potential surfaces for (a) Manα1–2Manα, (b) Manα1–3Manα, and (c) Manα1–6Manα computed by MNDO methods. Experimental (NOE-derived) and theoretical ϕ and ψ values are denoted Δ and +, respectively. The experimental values are derived from the corresponding linkages in structure VI of Figure 7. In panels, contour levels indicating energy bounds for 70%, 95%, and 99% of the molecules are shown.

experimentally derived (NOE) conformations are a far better measure of the true solution conformation, provided the possibility of conformational averaging can be ruled out. For the majority of glycosidic linkages, the NOE-derived conformation lies close to the computed energy minimum, which itself is within a single, deep potential well (Figure 6) (Homans et al., 1987b,c). This suggests very strongly that the NMR-derived conformation is unique. In addition, the depth of the well can be used to estimate the magnitude of torsional oscillations about the NMR-derived conformation. This leads directly to the question of oligosaccharide dynamics.

Oligosaccharide Dynamics

Since the NOE data can be interpreted in terms of a single, low-energy conformer in most glycosidic linkages studied to date, any attempt to intepret these data in terms of motional averaging is at least ambiguous. As mentioned above, molecular orbital calculations suggest that motional averaging is restricted to narrow torsional oscillations in the majority of linkages. Notable exceptions are Manα1–6Manα and Manα1–6Manβ linkages and probably α1–6 linkages in general. In these cases, the potential well is relatively shallow (Homans et al., 1987c). Furthermore, the NMR-derived conformation, when interpreted in terms of a single conformer, is often not in good agreement with the computed energy minimum, although it is still a low-energy structure (Figure 6). These considerations suggest that the α1–6 linkage is disordered. Recently, the magnitude of torsional oscillations in linkages such as Manα1–3Manβ and Manα1–6Manα has been estimated by simulation of molecular dynamics trajectories (Homans et al., 1987c). Over a period of 10 pS, Manα1–3Manβ was found to undergo torsional oscillations of around ±20°. In contrast, Manα1–6Manα was found to undergo much larger torsional oscillations. In the case of the latter, the NOE data could be interpreted in terms of an average over many different conformers, the bounds of which were well predicted from the depth of the potential well calculated from MNDO methods. This average is not a linear average, but an average in "NOE space", which is a vector

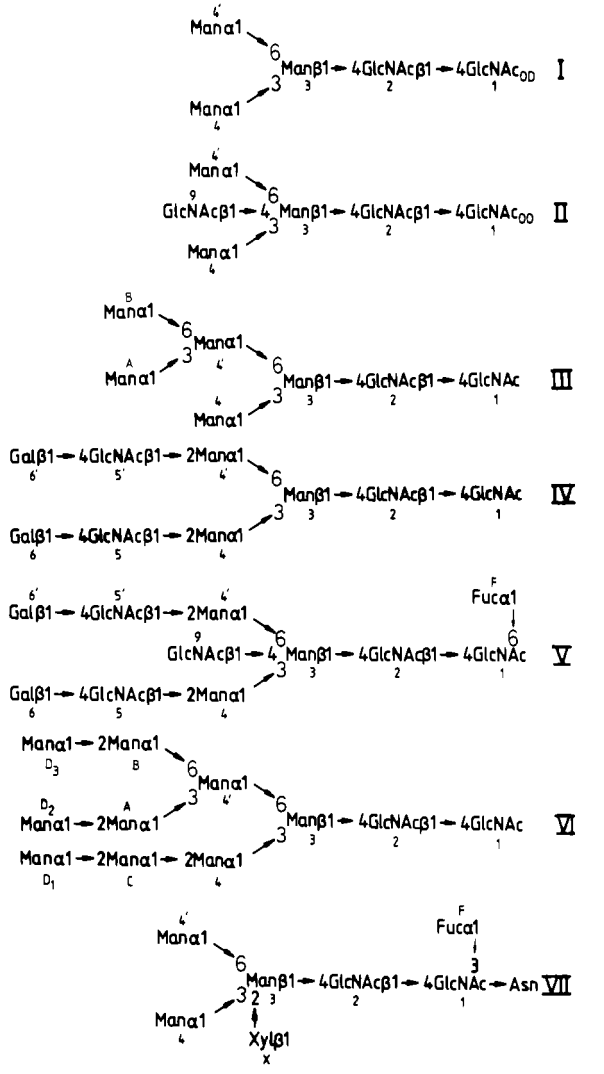

FIGURE 7: Representative structures of N-linked oligosaccharides studied to date. The roman numerals correspond with those used to describe the structures in the text.

space in which all internuclear distances are scaled to the inverse sixth power. The experimental "NOE average" structure may thus bear little relation to the true average.

Oligosaccharide Conformations

By use of the above NMR methods, together with molecular orbital calculations and molecular dynamics simulation, the three-dimensional structures of a variety of different yet structurally related oligosaccharides have been determined. A representative selection of these oligosaccharides is shown in Figure 7. We describe below only the pertinent aspects of the solution conformations of these oligosaccharides. Additional details can be found elsewhere (Homans et al., 1986, 1987b,c).

Since each oligosaccharide in Figure 7 contains the common pentasaccharide core structure I (see also Figure 1), it is logical to summarize the conformational properties of this at the outset. This structure contains regions of joint "rigidity/flexibility". The rigidity is associated with the Manα1–3Manβ1–4GlcNAcβ1–4GlcNAc fragment (Homans et al., 1982; Brisson & Carver, 1983; Homans et al., 1986, 1987b,c). From molecular dynamics studies the term rigidity implies narrow torsional oscillations (±20°) about a defined conformer. In contrast, the Manα1–6Manα linkage is, by the same criteria, disordered. This disorder, or flexibility is

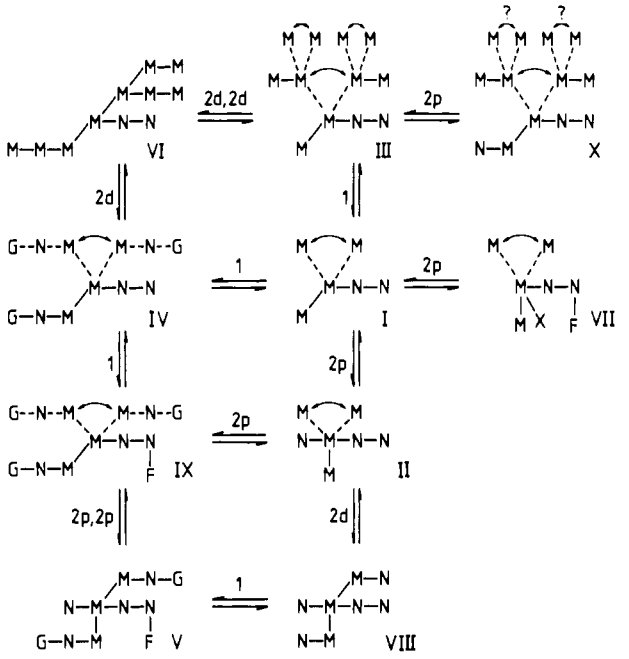

FIGURE 8: Conformational transitions between the oligosaccharides illustrated. This is an updated version of Figure 8 in Homans et al. (1986). Each monosaccharide residue is defined by a single letter code as follows: M, mannose; G, galactose; N, *N*-acetylglucosamine; F, fucose; X, xylose. A type 1 primary sequence change has no effect upon the primary sequence. In contrast, a type 2 sequence change alters the overall solution conformation, and this conformational transition may be either proximal (2p) or distal (2d) to the site of the primary sequence change. Dotted lines and double-headed arrows are used where more than one rotamer is found about the C5–C6 bond of Manβ of the pentasaccharide core. The roman numerals correspond to the structures shown in Figure 7. In addition, structure VIII is GlcNAcβ1–2Manα1–6(GlcNAcβ1–4)(GlcNAcβ1–2Manα1–3)-Manβ1–4GlcNAcβ1–4GlcNAc, studied by Strecker et al. (1977); structure IX is Galβ1–4GlcNAcβ1–2Manα1–6(Galβ1–4GlcNAcβ1–2Manα1–3)Manβ1–4GlcNAcβ1–4(Fucα1–6)GlcNAc and structure X is Manα1–6(Manα1–3)Manα1–6(GlcNAcβ1–2Manα1–3)Manβ1–4GlcNAcβ1–4GlcNAc, both of which have been investigated by Brisson and Carver (1983a,b). In the case of structure X, no experimental evidence is available to suggest that averaging occurs about the Manα1–6Manα linkage, but this is implied from chemical shift data (Homans et al., 1986).

twofold. First, there are two rotamer distributions about the C5–C6 bond of Man-3, which are approximately equally populated. These are termed the $\omega = 180°$ rotamer and the $\omega = -60°$ rotamer. Second, molecular dynamics simulations, together with NOE data, show that Manα1–6Manβ glycosidic linkages are disordered, unless they are restrained by long-range interactions, dispersive or otherwise (see below).

The influence of the so-called "bisecting" GlcNAc (βGlcNAc in 1–4 linkage to Man-3) upon the conformation of the pentasaccharide core can be found from similar investigations on compound II. Here, the rotamer distributions about the C5–C6 bond of Man-3 are again found to exist with $\omega = 180°$ and $\omega = -60°$. In contrast, the solution conformation about the Manα1–3Manβ linkage is found to undergo a small change, which is accompanied by characteristic chemical shift perturbations. This change is very subtle, and it is necessary to use a specialized technique, termed rotating frame Overhauser effect spectroscopy (ROESY), to demonstrate the conformational variation (Davis & Bax, 1985; Homans et al., 1987b). It is, however, well predicted from molecular orbital calculations on the relevant model compounds (Homans et al., 1987b).

While the addition of the bisecting GlcNAc causes a conformational change in the oligosaccharide, this result is by no means general. For example, a comparison between structure I and structure IV shows no discernible differences between the two with regard to the orientation of the Manα1–3Manβ linkage or the rotamer distributions about C5–C6 on Manβ. However, the addition of a bisecting GlcNAc to compound IV not only causes the above-mentioned conformational change in the Manα1–3Manβ linkage but also causes the value of ω to be restricted to 180°. The net result is that the α1–6 antenna maps a conformational space which is in the vicinity of the core region and is thus "folded back" toward it. A related phenomenon occurs in oligomannose-type oligosaccharides. In structure III, ω again exists in an equilibrium between 180° and −60°. In contrast, ω is restricted to 180° in structure VI. A fundamental difference between this conformational change and that described for structures IV and V is that in the latter the bisecting GlcNAc is *proximal* to the site of the conformational change, whereas the site of glycosylation in VI is distal to the site of the conformational change.

We have attempted to categorize these "conformational transitions" in terms of the primary sequence change causing them. When the addition or deletion of a monosaccharide does not cause a conformational transition, this is defined as a type 1 change. Conversely, if a conformational transition occurs, this is defined as a type 2 change. From the above discussion it is clear that the latter can be subcategorized according to whether the conformational transition is proximal or distal to the site of the addition or deletion. The former is termed a type 2p sequence change, whereas the latter is termed a type 2d sequence change. With this in mind, the various relationships between the overall conformations of the structures that have been investigated to date can be represented by an interconnecting network (Figure 8). This representation is useful in summarizing the results and also illustrates some of the more subtle aspects of the dependence of three-dimensional structure upon primary sequence. For example, it is seen that the presence of *both* a bisecting GlcNAc residue *and* a GlcNAc residue in the α1–6 antenna is required to restrict the value of ω to 180° in structure VIII. Such cooperativity suggests that α1–6 antenna–core interactions may have an important influence upon "restrained" structures. This is consistent with recent observations which demonstrated that in structure VI the α1–6 antenna *as a whole* was restrained and occupied a conformational space in the vicinity of the core (Figure 9). However, it was not possible from the available data to determine whether any arm–core interactions exist (Homans et al., 1987c).

Implications of Oligosaccharide Conformation to Biological Recognition

The results summarized in Figure 8 have three immediate implications. First, they demonstrate that oligosaccharides exist in solution with regions of defined three-dimensional structure. In general, oligosaccharides represent a class of molecules whose members are structurally very similar and would map similar regions of an *n*-dimensional conformational space if the number or degrees of freedom (n) were large. Thus, conformational rigidity is required to maintain specificity in recognition.

The second implication derives from the fact that the overall conformation is modulated by certain "key" residues, whereas other residues have no great influence. One of the greatest difficulties with the recognition hypothesis is the phenomenon of microheterogeneity, that is, the occurrence of a manifold of oligosaccharides of dissimilar (but often related) primary

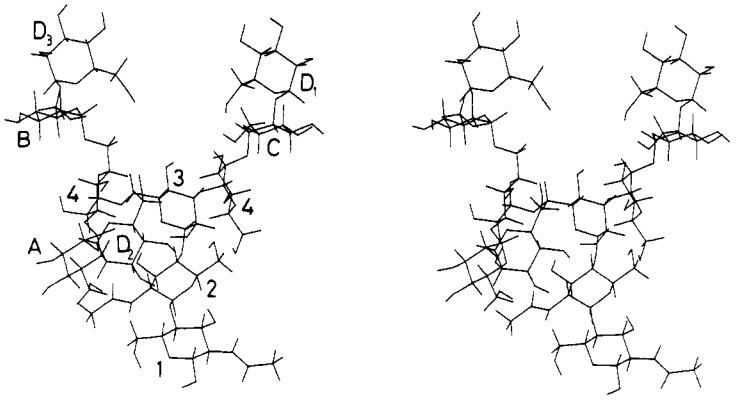

FIGURE 9: Stereoview of structure VI (see Figure 7 for nomenclature). Note that the Manα1–3Manα branch of the α1–6 antenna maps a conformational space that is close to the chitobiose core.

sequence at a given glycosylation site (i.e., glycoforms of a polypeptide). Although it has been suggested that microheterogeneity serves to diversify protein function, its extent is at first difficult to rationalize in terms of molecular recognition. For example, the oligosaccharides at Asn-297 in the constant region of the heavy chains of immunoglobulin G form a set of over 30 discrete structures (Rademacher et al., 1985). This diversity is, however, at the level of the primary sequence. The modulation of key residues serves to restrict the number of discrete overall conformations to a more limited subset, and it is these conformations that may be relevant to biological activity. Thus, individual members of a given set of primary sequences could all be recognized by a single specific receptor, albeit with different affinity, provided that the presence or absence of a key residue is common to all members of that set. This hypothesis certainly explains the binding specificity of certain lectins of plant origin but avoids the issue of the *purpose* of oligosaccharide microheterogeneity.

The third implication concerns oligosaccharide dynamics. Since the biantennary α1–3/α1–6 pentasaccharide core is conserved, the mobility of the α1–6 antenna as a whole confers the potential for joint rigidity/flexibility to all Asn-linked oligosaccharides (and incidentally to a number of glycolipids). The majority of structures in Figure 8 actually possess this duality. An obvious question is whether this is functionally relevant. As suggested by Montreuil (1984), one possibility lies in the interaction of an oligosaccharide with the attached protein. A good example of this is the Fc fragment of immunoglobulin G (Diesenhofer, 1981; Sutton & Phillips, 1983). In the crystal structure, the α1–6 antenna of the complex type oligosaccharide (which is of course a superposition of many structures) attached at Asn-297 interacts extensively with the protein. In contrast, the α1–3 antenna is oriented toward the oligosaccharide on the symmetry-related domain. The results of studies on the Manα1–6Manβ glycosidic linkage indicate that it is in a low-energy configuration (Homans et al., 1987c), and therefore the interaction of the α1–6 antenna with the protein is not unfavorable on this basis. Whether this is a general phenomenon remains to be seen, but the increased susceptibility to proteases of many glycoproteins after glycosidase treatment strongly suggests that oligosaccharide–protein interactions such as those found in the Fc fragment are not unique. Thus, rather than considering the oligosaccharide in isolation, we must also consider the possibility that an oligosaccharide–protein surface might act as a ligand.

REFERENCES

Bax, A., & Freeman, R. (1981) *J. Magn. Reson. 44*, 542.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 63*, 207.

Bock, K., Arnarp, J., & Lonngren, J. (1982) *Eur. J. Biochem. 129*, 171.

Boyd, J., & Bazzo, R. (1987) *J. Magn. Reson.* (in press).

Brisson, J.-R., & Carver, J. P. (1983a) *Biochemistry 22*, 3671.

Brisson, J.-R., & Carver, J. P. (1983b) *Biochemistry 22*, 3680.

Calvo, F. O., & Ryan, R. J. (1985) *Biochemistry 24*, 1953.

Clore, G. M., & Gronenborn, A. M. (1985) *J. Magn. Reson. 61*, 158.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820.

Deisenhofer, J. (1981) *Biochemistry 20*, 2361.

Dennis, J. W., & Laferté, S. (1985) *Cancer Res. 45*, 6034.

Dewar, M. J. S., & Thiel, W. (1977) *J. Am. Chem. Soc. 99*, 4899.

Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731.

Ferguson, M. A. J., Homans, S. W., Dwek, R. A., & Rademacher, T. W. (1987) *Science (Washington, D.C.)* (submitted for publication).

Homans, S. W., Dwek, R. A., Fernandes, D. L., & Rademacher, T. W. (1982) *FEBS Lett. 150*, 503.

Homans, S. W., Dwek, R. A., Fernandes, D. L., & Rademacher, T. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6286.

Homans, S. W., Dwek, R. A., Boyd, J., Mahmoudian, M., Richards, W. G., & Rademacher, T. W. (1986) *Biochemistry 25*, 6342.

Homans, S. W., Dwek, R. A., Boyd, J., Soffe, N., & Rademacher, T. W. (1987a) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1202.

Homans, S. W., Dwek, R. A., & Rademacher, T. W. (1987b) *Biochemistry 26*, 6553–6560.

Homans, S. W., Pastore, A., Dwek, R. A., & Rademacher, T. W. (1987c) *Biochemistry* (in press).

Hudgin, R. L., Pricer, W. E., Jr., Ashwell, G., Stockert, R. J., & Morell, A. G. (1974) *J. Biol. Chem. 249*, 5536.

Lindon, J. C., Vinter, J. G., Lifely, M., & Moreno, C. (1984) *Carbohydr. Res. 133*, 59.

Montreuil, J. (1984) *Biol. Cell 51*, 115.

Morell, A. G., Irvine, R. A., Sternlieb, I., Scheinberg, I. H., & Ashwell, G. (1968) *J. Biol. Chem. 243*, 155.

Muller, L., & Ernst, R. R. (1980) *Mol. Phys. 38*, 963.

Paulsen, H., Peters, T., Sinnwell, V., Lebuhn, R., & Meyer, B. (1984) *Liebigs Ann. Chem.*, 951.

Paulsen, H., Peters, T., Sinnwell, V., Heume, M., & Meyer, B. (1986) *Carbohydr. Res. 156*, 87.

Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800.

Prieels, J. P., Pizzo, S. V., Glasgow, L. R., Paulson, J. C., & Hill, R. L. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2215.

Rademacher, T. W., Homans, S. W., Parekh, R. B., & Dwek, R. A. (1985) *Biochem. Soc. Symp. 51*, 131.
Sharon, N., & Lis, H. (1982) *Proteins (3rd Ed.)*, 1–144.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286.
Strecker, G., Herlant-Peers, M.-C., Fournet, B., Montreuil, J., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., & Farriaux, J. P. (1977) *Eur. J. Biochem. 81*, 165.
Sutton, B. J., & Phillips, D. C. (1983) *Biochem. Soc. Trans. 11*, 130.
Vliegenthart, J. F. G., Dorland, L., & van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 209.
Wüthrich, K., Wider, G., Wagner, G., & Braun, W. (1982) *J. Mol. Biol. 155*, 311.

Accelerated Publications

Generation and Statistical Mechanical Modeling of Z-DNA in the Mouse Metallothionein I Promoter†

Daniel Visentin and Calvin B. Harley*
Medical Sciences Program and Department of Biochemistry, McMaster University, Hamilton, Ontario, Canada L8N 3Z5
Received July 9, 1987; Revised Manuscript Received August 12, 1987

ABSTRACT: The mouse metallothionein I promoter has a potential Z-DNA sequence 86–102 bases upstream from the transcription start. The alternating purine/pyrimidine nature of this sequence, d-(GCGCGTG*AC*TATGCGTG), is interrupted by a central d(AC) dinucleotide. We reversed the d(AC) to d(CA) by site-directed mutagenesis and subcloned the wild-type and mutated 218 base pair (bp) *Sst*I–*Bgl*II fragment containing this region into pUC12. By two-dimensional chloroquine gel analysis, the mutant promoter clearly underwent a B–Z transition with a resulting loss of at least five twists. Purines within this sequence exhibited diethyl pyrocarbonate (DEP) sensitivity, which extended in both the 5′ and 3′ direction encompassing a region of approximately 32 base pairs. When subjected to sufficient torsional stress, the wild-type sequence showed weak evidence of a transition on chloroquine gels and clear DEP sensitivity with a similar, yet distinct pattern. Statistical mechanical modeling of the chloroquine gel analysis demonstrated that the average free energy of propagation for the mutant sequence (0.7–0.9 kcal mol^{-1} bp^{-1}) was approximately twice that for $d(CG)_n$ sequences and that the sequence d(GACGCGGGGCGCGTGCA-TATGCGTGG) forms the core Z-DNA region.

Left-handed Z-DNA and its dynamics are well characterized for synthetic oligonucleotides in vitro [reviewed in Rich et al. (1984)]. Alternating purine/pyrimidine (pur/pyr)[1] sequences most readily form Z-DNA in a syn/anti conformation, respectively (Haschemeyer & Rich, 1967; Wang et al., 1979, 1984), but only two of the three redundant alternating pur/pyr sequences [$d(CG)_n$ and $d(CA)_n$] adopt Z-conformation when subjected to the torsional stress of negative superhelicity (Singleton et al., 1982; Peck et al., 1982; Peck & Wang, 1983; Haniford & Pulleyblank, 1983; Wang, 1984). $d(AT)_n$ forms Z-DNA under certain conditions (Patel & Kozlowski, 1985; Adam et al., 1986; McLean et al., 1986) but preferentially adopts a cruciform structure under superhelical stress (Greaves et al., 1985), as does $d(CATG)_{10}$ (Naylor et al., 1986). Although Z-DNA formation is enhanced by solvent conditions, base modification, and protein binding (Pohl & Jovin, 1972; Wang et al., 1984; Nordheim et al., 1982; Azorin & Rich, 1985), it is clearly sequence dependent. The presence of d(A/T) base pairs or interruptions of pur/pyr alternation may impose a substantial energetic cost to Z-DNA formation (Wang et al., 1984; Patel & Kozlowski, 1985; McLean et al., 1986; Ellison et al., 1986).

Indirect evidence suggests that Z-DNA occurs naturally (Nordheim & Rich, 1983; Rich et al., 1984) and may be involved in recombination (Slightom et al., 1980; Kmiec & Holloman, 1984), DNA packaging (Miller et al., 1985), and gene regulation (Nordheim & Rich, 1983; Hipskind & Clarkson, 1983). However, sequence constraints on the B–Z transition of natural DNA are poorly understood. The mouse metallothionein I (mMT-I) promoter has potential Z-DNA sequences adjacent to regulatory regions. Similar observations were made for the human metallothionein IIA promoter (Karin et al., 1984). We investigated this region in the

† This research was supported in part by the Medical Research Council of Canada and the Natural Sciences and Engineering Research Council of Canada during the tenure of an MRC Scholarship (C.B.H.) and a Gerontological Research Council of Ontario Advanced Student Bursary held within the Medical Sciences Growth and Development Program (D.V.).

[1] Abbreviations: pur/pyr, purine/pyrimidine; mMT-I, mouse metallothionein I; 2-D, two dimensional; bp, base pair(s); TBE, Tris–borate–EDTA; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; DEP, diethyl pyrocarbonate.

0006-2960/87/0426-6578$01.50/0