Journal of Biomolecular NMR, **21:** 127–139, 2001.
KLUWER/ESCOM

Determining binding sites in protein–nucleic acid complexes by cross-saturation

Andrew N. Lane[a,*], Geoff Kelly[b], Andres Ramos[a] & Thomas A. Frenkiel[b]
[a]*Division of Molecular Structure and* [b]*MRC Biomedical NMR Centre, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.*

Received 25 June 2001; Accepted 7 August 2001

Key words: cross-saturation, Mbp1-DNA interaction, spin diffusion

Abstract

Cross-saturation experiments have been shown to give accurate information regarding the interacting surfaces in protein-protein and protein-RNA complexes. The rate of magnetization transfer depends on a number of factors including geometry, spin topology, and effective correlation times. To assess the influence of these variables on such experiments, and to determine the range of applicability of the technique, we have simulated the time-course of magnetization transfer across the interface in a variety of protein-nucleic acid complexes (434 Cro, SRY, MetJ and U1A), each having a different binding geometry. The simulations have been carried out primarily to provide information about the experimentally accessible targets for selective saturation, such as the anomeric protons and the imino protons of the nucleic acid. Saturation of either of these groups of signals leads to partial excitation throughout the nucleic acid molecule, and the resulting transfer of saturation to the labelled protein moiety can be readily detected by the reduction in intensity of particular peaks in the HSQC spectrum of the protein. The simulations show that information can be obtained about the residues in contact with the nucleic acid without any need for deuteration. Experimental cross-saturation data have been obtained from the Mbp1-DNA complex and interpreted in conjunction with the simulations to map out the binding surface in detail.

Abbreviations: Mbp1, Mlu1 cell cycle box binding protein.

Introduction

Identifying the contact surfaces of interacting molecules is important for several reasons, such as providing a framework for analyzing mutation experiments and guiding docking calculations. It may also be the first step in giving a structural basis for the function of a complex. Methods of defining binding surfaces that are more straightforward than a full structure determination are therefore of considerable value. Methods have been proposed based on chemical shift perturbations (Foster et al., 1998; Williamson et al., 1997) and amide hydrogen exchange (Paterson et al., 1990), but the precision of these techniques is low. Recently, a cross-saturation experiment was described to map the interaction surfaces in protein-protein complexes (Takahashi et al., 2000). Here a non-deuterated ^{14}N protein, the target, was complexed with a perdeuterated ^{15}N-enriched protein, the reporter. Saturation of the aliphatic region of the protiated target protein results in a general loss of magnetization throughout this component, aided by spin diffusion. Transfer of saturation from protons of the target protein to the amide protons of the reporter protein results in decreased intensity in the ^{15}N-^{1}H HSQC spectrum. The largest effects were observed for residues in the interface, as spin diffusion through the reporter protein is much attenuated by the deuteration. The principle was recently extended to mapping protein-RNA interactions without the need for deuteration (Ramos et al., 2000), taking advantage of the fact that RNA has a number of resonances (mainly ribose anomeric

*To whom correspondence should be addressed. E-mail: alane@nimr.mrc.ac.uk

protons) which are not generally overlapped by signals of the protein. These signals can be used to selectively excite the RNA even in a fully protiated complex, and again transfer of saturation occurs from the RNA to the protein. It was shown experimentally that the cross-saturation experiment can be used to determine the interaction surface in protein–RNA complexes (Ramos et al., 2000).

Because the influence of spin-diffusion in the fully protiated complex was surprisingly small (Ramos et al., 2000), we have carried out simulations of the saturation transfer in complexes of known structure in order to establish the limits of the methodology, and to extend it to protein–DNA complexes. The simulations show that details of protein–nucleic acid interfaces can be defined, as well as the relative orientations of the interacting components, without the need for deuterated protein. Different measurement and excitation modes of the unlabelled components can provide complementary information that specifically pick out the interaction interface.

We have simulated the saturation transfer between DNA and a variety of proteins in complexes that have very different architectures and modes of binding. 434 Cro is an archetypal helix-turn-helix protein that binds to the major groove of DNA (Mondragon and Harrison, 1991). SRY is a mainly helical protein with an extended N-terminal strand that binds to the minor groove side of the DNA and distorts it by intercalating Ile 13 between two bases (Werner et al., 1995). MetJ is a homodimer in which the primary DNA binding surface is a short antiparallel beta sheet formed from strands on the two subunits (Somers and Phillips, 1992). In addition, we have simulated the saturation transfer in the complex between the protein U1A (1–102) and a 39 nucleotide RNA, the structure of which was solved by NMR (Allain et al., 1997). This provides an example of an RNA-protein complex and involves interactions with a different nucleic acid geometry: The RNA forms two duplex stems connected by a long (7 nucleotide) loop on the outer surface and a short loop on the inner surface, and the protein makes contacts primarily to the outer loop, with a few contacts to the tops of the stem. We precede these simulations on protein–nucleic acid complexes with some simulations which demonstrate the saturation transfer within B DNA alone, using as an example the DNA component of the above 434 Cro-DNA complex.

The simulations have been carried out for proteins labelled with ^{15}N only and proteins labelled with both ^{15}N and ^{13}C. We have examined the time dependence of the magnetization transfer through ^{13}C-labelled side chains, for correlation times up to 30 ns, and for semi-selective saturation of the imino protons of the DNA, the anomeric protons of the DNA, and also for selective excitation of individual anomeric protons. We have also simulated saturation transfer from Cyt H5, because these protons may overlap some anomeric protons. The effects of anisotropic molecular rotation and internal motion of side chains have been assessed using simplified motional models.

The simulations have been used to design and interpret cross-saturation experiments with the DNA-binding domain of the yeast cell-cycle transcription factor Mbp1 (McIntosh et al., 1991; Lowndes et al., 1992). The structure of the *N*-terminal 124 residue DNA-binding domain has been determined by X-ray crystallography (Taylor et al., 1997; Xu et al, 1997), and the DNA binding site has been located by chemical shift mapping (Taylor et al., 2000) and ^{15}N relaxation (McIntosh et al., 2000). The cross-saturation experiments permit a more precise determination of the DNA-binding surface of the protein.

Materials and methods

Materials

Mbp1 and its complex with DNA were prepared as previously described (Taylor et al., 2000). The Mbp1 protein was analyzed for purity by SDS PAGE and for DNA-binding activity using a DNA band-shift binding assay with the synthetic 14 base-pair DNA duplex sequence: d(CTTACGCGTCATTG).d(CAATGACGCG TAAG) (consensus binding site underlined) as previously described (Beckman et al., 1993). The two DNA strands were synthesised and purified by reverse phase HPLC by Oswel Research Products Ltd. (Southampton, U.K.).

NMR spectroscopy

NMR spectra were recorded at 11.75 T and 14.1 T on Varian Unityplus and Varian Inova spectrometers, respectively, at 15 °C. Cross-saturation experiments were carried out essentially as described by Ramos et al. (2000), except that simple continuous wave irradiation was used in place of a band-selective pulse in those experiments where only a single resonance was to be saturated.

Simulations

Co-ordinates of the various protein–DNA complexes were taken from the protein database. Water molecules were not included in these simulations as they are incomplete in the deposited files, or absent in the case of NMR structures. A co-ordinate file for B DNA alone was extracted from the file for the 434 Cro complex by eliminating the co-ordinates of all the other atoms. Hydrogen atoms were added, and the system allowed to relax under the influence of the Amber forcefield with 500 cycles of conjugate gradient energy minimization. This is essential to regularise all structures in the same way and to eliminate any bad hydrogen-hydrogen contacts that would otherwise lead to erroneously large proton relaxation rate constants. Magnetization evolution was calculated using the program Prophet (A.N. Lane, unpublished data), in which the hydrogen co-ordinates are used to calculate the relaxation matrix, R, using the following definitions for the cross-relaxation rate constant (σ) and auto-relaxation rate constant (ρ):

$$\sigma_{ij} = (\alpha/r_{ij}^6)S_{ij}^2[6J(2\omega) - J(0)], \quad (1)$$

$$\rho_i = \sum(\alpha/r_{ij}^6)S_{ij}^2[J(0) + 3J(\omega) + 6J(2\omega)] \quad (2)$$

where α is a function of nuclear constants (56.92 s^{-1} Å^6 ns^{-1} for proton-proton interactions), $J(\omega) = \tau/(1 + \omega^2\tau^2)$. Here, internal motion is assumed to be sufficiently rapid such that spectral density functions are simply scaled by the generalised order parameter S^2. The generalised order parameter S^2 would be equal to unity for a rigid molecule. To model internal motion, the order parameter for the vector between protons H_i and H_j was calculated as $(S_i^2S_j^2)^{0.5}$, where S_i^2, $S_j^2$are the order parameters for individual N-H or C-H vectors. S^2 was taken as unity for NH and 0.9, 0.8, 0.7, 0.6 and 0.5 for Hα, Hβ, Hγ, Hδ and Hε, respectively. For aromatic rings S^2 was assigned the value of unity and in some simulations a value of 0.5. Additional terms were included in Equation 2 to describe auto-relaxation by attached ^{13}C or ^{15}N where appropriate. Simulations were carried out for a spectrometer frequency of 600 MHz. Order parameters involving methyl groups were calculated based on the Tropp (1980) model, as previously described (Lane, 1990). Internal relaxation of the methyl group was calculated assuming fast rotation as previously described (Lane and Fulcher, 1995).

The evolution of the magnetization of all the protons for a chosen initial starting state (e.g., saturation of all NH or all H1′ protons) was calculated by numerical integration of the system of Solomon equations, using a fourth-order Runge–Kutta method as previously described for the program NUCFIT (Lane, 1990). The integration step-size was set to $1/(5\rho_{max})$, where ρ_{max} is the largest spin-lattice relaxation rate constant in the system. Continuous saturation of one or more spins was simulated by resetting the magnetization of these spins to zero at each stage of the numerical integration.

Motional anisotropy was treated using the calculated moments of inertia about three orthogonal axes. These were calculated for all orientations about the z and x axes in 1 degree steps, according to the formula:

$$I = \Sigma m_i r_i^2 \quad (3)$$

where m_i is the mass of atom i and r_i is the distance from the atom to the axis in question. This provides the principal axes of the moment of inertia tensor. The molecule was oriented with the smallest value of I (longest axis of the molecule) along the z axis. The anisotropy was calculated as $2I_z/(I_x+I_y)$. Similarly, the radius of gyration, Rg, was calculated using Equation 3 but with r_i defined as the position of atom i from the centre of mass.

As the complexes are nearly axially symmetric, the simplified Woessner model (1962) was used to calculate the orientational components of the spectral densities, viz:

$$\begin{aligned} J(\omega) &= a_1J(\omega, \tau_1) + a_2J(\omega, \tau_2) \\ &\quad + (1 - a_1 - a_2)J(\omega, \tau_3), \end{aligned} \quad (4)$$

where $a_1 = 0.25(3\cos^2\beta - 1)^2$, $a_2 = 3\cos^2\beta\sin^2\beta$.

β is the angle between the H-H vector and the z axis. τ_i are rotational correlation times:

$$\begin{aligned} \tau_1 &= \tau_L, \tau_2 = 6\tau_L\tau_s/(5\tau_s + \tau_L), \\ \tau_3 &= 3\tau_L\tau_s/(2\tau_L + \tau_s), \end{aligned} \quad (5)$$

where τ_L and τ_s are correlation times for rotation of the long and short axes, respectively. These correlation times were calculated using the Perrin equations (Cantor and Schimmel, 1980) and the given value of the isotropic rotational correlation time τ_0.

An operational definition of the interface in the complexes was adopted: protein residues that contain at least one atom within 3.0 Å of an atom in the nucleic acid molecule were taken to be interfacial. Simulation results are presented as NOEs:

$$NOE = (M_{sat} - M_0)/M_0, \quad (6)$$

where M_{sat} is the magnetization of a particular proton in the presence of the perturbation and M_0 is the equilibrium magnetization of the proton.

Results

Time courses of the changes in magnetization of complete spin systems were simulated as described in Materials and methods. Table 1 summarizes salient features of the test systems simulated.

Saturation transfer within the DNA duplex from the 434Cro complex

Figure 1A shows the saturation transfer in the DNA alone following saturation of the H1′ or imino protons for 1 s. In general, the spin-diffusion is more extensive starting from H1′ than from NH. This is because with H1′ saturation every nucleotide is directly excited, and also the H1′ protons are close to more protons which interact with the bases (and therefore transfer saturation from the minor groove to the major groove). In contrast, imino protons are sparser (only one per base pair), and they are located in the center of the duplex, close to the helix axis, and interact directly only with amino protons of A and C and to a lesser extent the H8 of A and H6, H5 of C (all in the major groove), and only GNH_2 and AH2 in the minor groove. (These amino protons are often involved in the direct readout hydrogen-bonding interactions.) These results suggest that in complexes H1′ saturation would be expected to show sizable effects for interfaces that involve either the major or the minor grooves, whereas the imino saturation experiment would be expected to pick out fewer contacts and to give smaller effects overall, especially for protein residues which bind in the minor groove.

Selective saturation of an individual anomeric proton can be achieved using a simple continuous wave pulse. The pattern of saturation transfer under this condition is much simpler, as only a small region of the DNA is excited to a significant degree, extending no further than the irradiated nucleotide, its 5′ and 3′ nearest neighbours on both strands, and the H-bonded nucleotide. Figure 1B shows the magnetization transfer on saturation of a single Thy 15 H1′ in the 434 Cro:DNA complex. The spread of saturation is rather localized, and is significant only for the excited nucleotide and its nearest neighbours. Hence, one would expect that in a complex, transfer to the protein would be mainly to those residues close to the irradiated nucleotide.

Saturation transfer within the 434Cro-DNA complex

The 434Cro protein is a classical helix-turn-helix DNA binding protein (Mondragon and Harrison, 1991). The recognition helix makes numerous contacts with exposed groups in the major groove of the DNA. In addition, there are several close contacts to the minor groove from the loop adjacent to the recognition helix on the same face of the protein. Figure 2 depicts a space-filling model of the 434Cro protein within the complex with OR1 solved at 2.5 Å (Mondragon and Harrison, 1991) showing the interface as defined by the distance criterion (see Methods). This comprises residues 10, 16–18, 27–32 and 42–44. However, the side chain of Arg 10 makes a salt bridge with a phosphate oxygen, and is not close to any DNA protons; such contacts will not be observed in cross-saturation experiments.

H1′ saturation

Simulation of the effects of irradiation of the anomeric protons of the DNA moiety for 1 s shows significant (3–12%) negative enhancements of the protein HSQC NH cross-peaks corresponding to residues 15–17, 26–33 and 42–44, with very small enhancements elsewhere (Figure 3A). As expected, Arg10 does not appear in the NOE plot (see above). Hence, the NOE profile is in excellent agreement with the residues identified as being in the interface using the simple distance criterion (shown also as filled squares in Figure 3A). The regions of strong enhancement correspond to the recognition helix (26–34) and the loop that is in close contact with the minor groove of the DNA (42–44). Figure 3B shows the time course of the build-up from the H1′ protons to the various protons of Phe 44. As expected, there is a substantial lag in the build-up curves that reflects first the spread of saturation from H1′ (minor groove) to base protons in contact with the protein side chains, and then subsequent spin-diffusion along the side chains toward the main chain. Thus, the ε and δ protons of the ring of Phe 44 become (partially) saturated first, followed by transfer to the more distant side chain and main chain protons. The final values of the magnetization also reflect the cross-relaxation pathways, with the ε and δ protons having the greatest (negative) enhancements.

Excitation of individual H1′ resonances results in saturation transfer to a more restricted set of protons

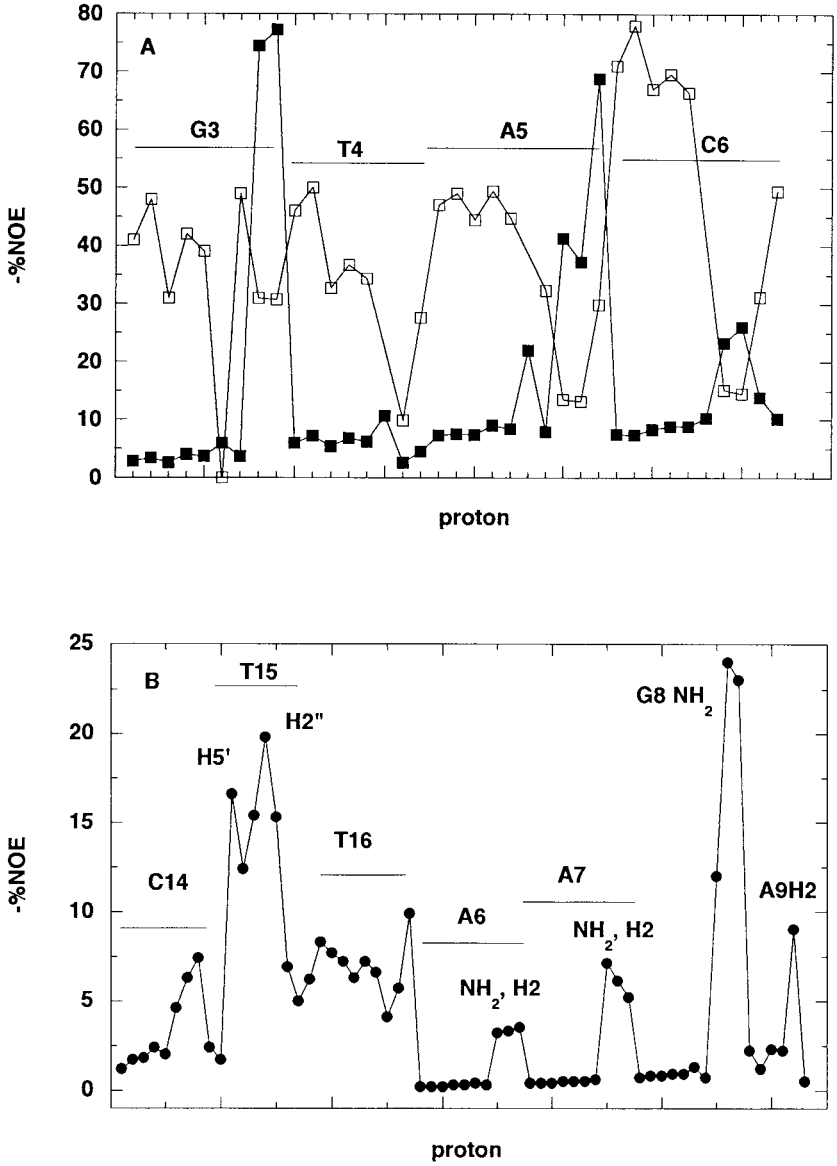


Figure 1. Saturation transfer within the DNA in the 434 Cro-DNA complex. Simulations were carried out for a rigid body (except methyl groups) using an isotropic rotational correlation time of 20 ns. The protons are shown sequentially for the residues shown in the order H5′/H5″, H4′, H3′, H2′, H2″, H1′, H8/H6, NH/NH_2. (A) Saturation transfer in the DNA after 1 s of irradiation of all H1′ (□) or all NH (■). (B) Saturation transfer on irradiation of H1′ of residue T15 for 1 s.

(cf. Figure 1). For example, selective excitation of Cyt 24 H1′ showed saturation transfer only to Gln 29 (not shown). Hence, this provides a simple way to map out both the protein residues and the nucleotides that are involved in the interface, provided that the anomeric protons have been assigned.

Cytosine H5 saturation

In practice, the resonances of the H1′ protons may overlap those of cytosine H5 protons, which are found in the major groove. Simulation of the effects of saturating all the H5 protons shows that significant saturation transfer occurs to the protons of the same residues as when saturating the H1′ or NH, but with different relative intensities. For example, the most intense NOE from H1′ saturation is to Phe 44, followed

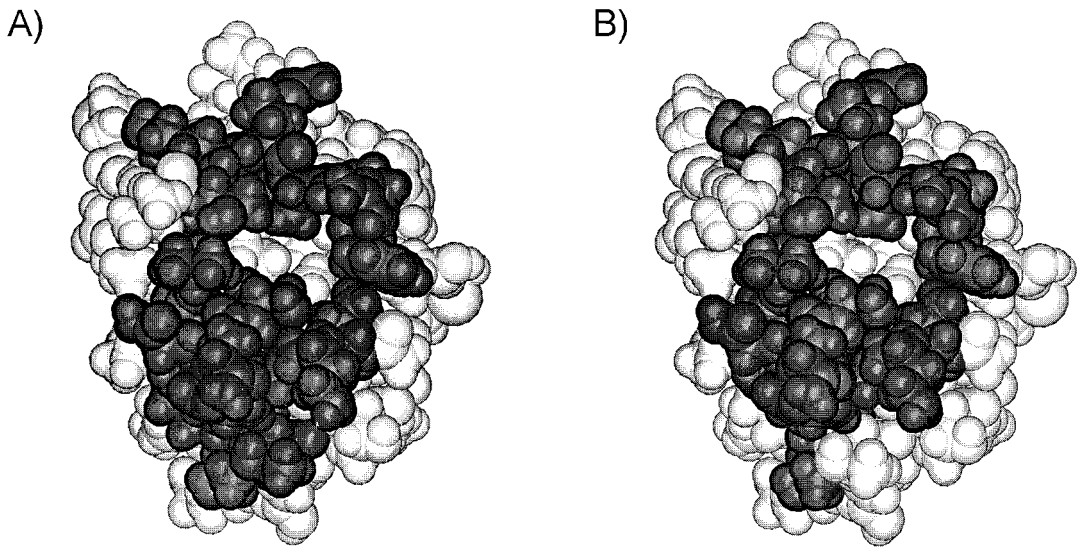


Figure 2. DNA-binding surface of 434Cro. (A) Residues comprising the DNA binding surface determined as at least one protein-DNA inter-atomic distance < 3 Å are shown in dark grey, and other residues in light grey. (B) Residues that show an NOE to H_N, Hα or Hβ of > 3% after saturating the H1′ for 1 s.

by Q32 > Q31 > K27 > V26 > Q17 > T16. For H5 saturation, the intensities were in the order V26 > Q32 > F44 > T16 (not shown).

Imino proton saturation

Saturation of the imino protons results in a similar pattern of NOEs as H1′ excitation, but with generally lower intensity (Figure 3C). However, there are some notable differences between the experiments. The order of intensities is Q29 > Q32 > Q17≈F44, which is more similar to H5 saturation than H1′ saturation. Protein residues 28, 29, 30 and 32 make contacts with major groove protons, including the H2′ and H8 of adenines and especially the amino protons of Cyt and Ade. As these amino protons are also in good dipolar contact with the imino protons (cf. Figure 1), this provides an efficient pathway of saturation transfer to the imino protons, and therefore residues 28 to 32 are relatively strongly represented in the imino cross-saturation experiment. The low intensity to residues Arg 43 and Phe 44 reflects the geometry of this interface; the side chain of Phe 44 makes close contact to phosphate oxygens, and to sugar atoms in the minor groove of the DNA, which are not in good dipolar contact with the DNA imino protons. Similarly, the low degree of saturation transfer to Thr 16 and Thr 17 can be attributed to their contacts being mainly with minor-groove protons.

Effect of internal motion and anisotropy

The core of most proteins is generally well defined, with only fast (sub-nanosecond) low amplitude dynamics, and in some cases slower events such as 180° flips of Phe and Tyr rings. Residues involved in making contacts with DNA are generally on the surface of the protein, and are therefore likely to show more complex and extensive dynamics. However, those residues that directly contact the DNA by electrostatic or H-bonding interactions may become rigidified in the complex. We have used a simplified model of rapid dynamics to estimate the influence of motion (see above). Because both cross and autorelaxation rate constants are scaled by the same factor, the steady-state magnetization is not affected by the internal motion. However, the dynamics as simulated slow down the rate of transfer of saturation, but the effect is generally small, and can be compensated for by increasing the saturation time. For example, the NOEs to Phe 44 after 1 s of irradiation (cf. Figure 3B) decreased only 1–2% using an order parameter of 0.5 compared with the fully rigid system.

Of the systems studied here, the 434Cro-DNA complex has the largest anisotropy according to the moments of inertia (Table 1). The rotational correlation times for an anisotropy of 2 will differ by about 50% for vectors parallel to or perpendicular to the principal axis, with the longer effective correlation time for vectors nearly parallel to the principal

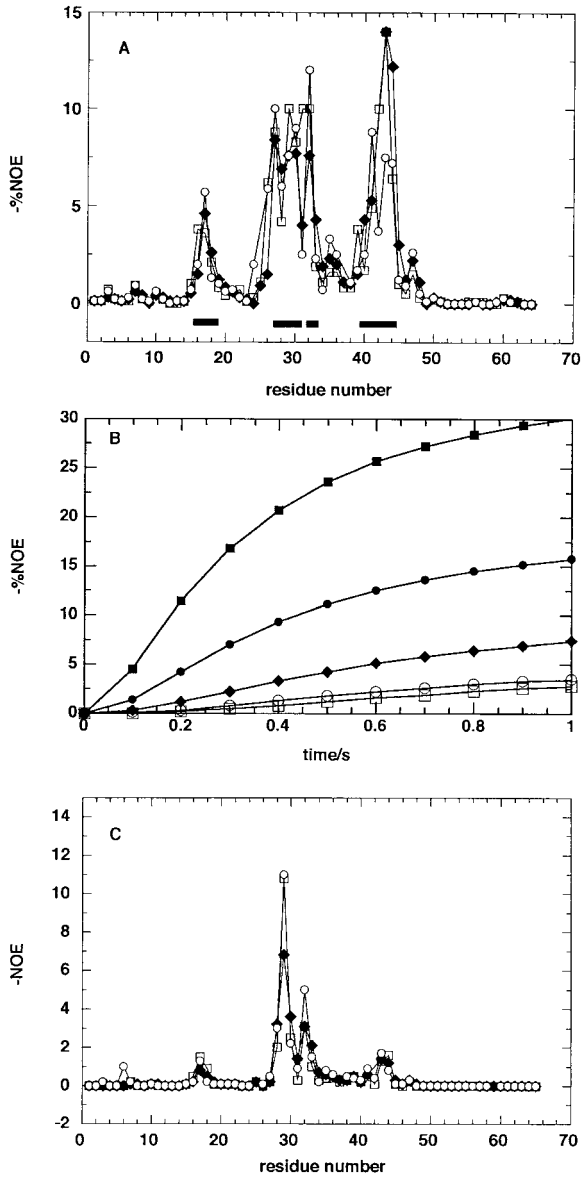


Figure 3. Cross-saturation in the 434 Cro-DNA complex. Simulations were carried out as described in the Methods section. For clarity, only one subunit is shown. (A) NOEs for saturation of the DNA H1′ for 1 s. (♦) H_N, (□) Hα, (○) Hβ. (■) residues with an atom within 3 Å of the DNA. (B) Time course of saturation transfer from H1′ to residue F44 (♦) H_N, (□) Hα, (○) Hβ, (●) Hδ1, (■) Hε1. (C) NOEs for saturation of the DNA imino protons for 1 s (♦) H_N, (□) Hα, (○) Hβ.

Table 1. Simulation conditions. All simulations were carried out at 600 MHz (^{1}H) for an H_2O solvent with ^{15}N-labelled protein. Rg is the radius of gyration. The anisotropy was calculated from the moments of inertia as described in Materials and methods

Complex	Id	MW	No. H	Rg/Å	Anisotropy
434cro:DNA	3cro	27100	1466	19.8	2.39
SRY:DNA	1hry	13700	732	15.4	1.89
metJ:DNA	1cma	30000	1762	18.7	1.26
U1A:RNA	1aud	21100	1095	18.6	2.20

axis. However, the main influence of this degree of anisotropy is comparatively small in the final degree of magnetization transfer (and see below).

Saturation transfer within the SRY-DNA complex

SRY is a mainly helical protein with an extended N-terminal strand that binds to the minor groove side of the DNA and distorts it by intercalating Ile 13 between two bases (Werner et al., 1995). The distance criterion defines the interface as comprising residues 4, 7, 8, 9, 10,12, 13, 17, 20, 33, 43, 62, 66, 73 and 74. There are almost no contacts within the major groove.

Figure 4 shows the results of cross-saturation simulations for irradiation of the imino protons and the H1′ resonances. Groups of substantial NOEs are observed for residues 4–14, 33–39, 62, 67 and 73–75, in agreement with the interface defined above and as shown as filled squares in Figure 4A. As expected, these residues are all on the same face of the protein molecule facing the DNA. Mostly, they do not correspond to the regular secondary structure elements. The absolute magnitude of the transfer is lower for saturation of the imino protons than for the H1′, and the details of the magnetization transfer also differ (Figures 4A and 4B).

The time course of the saturation transfer to H_N of the intercalating residue Ile 13 on saturation of the H1′ showed a significant lag, reaching about 8% at 1 s. This is because it is the methyl groups of Ile 13 that are in closest contact with an AdeH2. The ring protons of Tyr 74 showed very large (direct) NOEs, reaching 50% for δH1 at 1 s, whereas the amide proton showed a long lag period (data not shown). Nevertheless, a substantial NOE was attained by 1 s of irradiation (Figure 4A). This behaviour arises because the ring is in close contact with the DNA, and magnetization

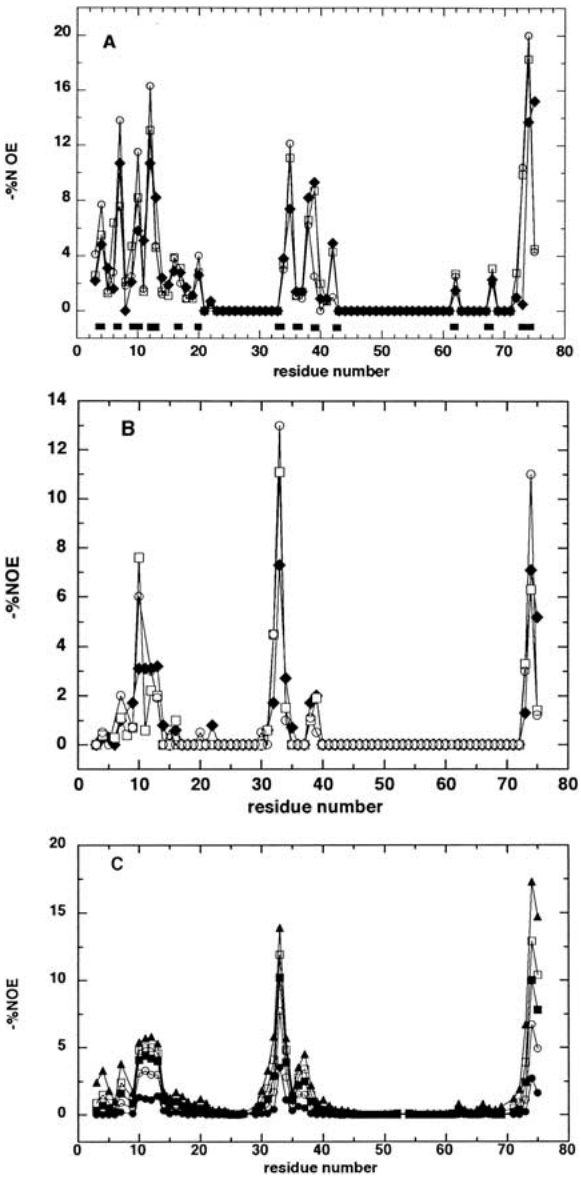


Figure 4. Cross-saturation in the SRY-DNA complex. Simulations were carried out using a correlation time of 10 ns. (A) NOEs for 1 s irradiation of H1′, (♦) H_N, (□) Hα, (○) Hβ. (■) residues with an atom within 3 Å of the DNA. (B) NOEs for 1 s irradiation of NH (♦) H_N, (□) Hα, (○) Hβ. (C) NOEs for 1 s irradiation of the imino protons, as a function of rotational correlation time. (●) 5 ns, (○) 10 ns, (■) 15 ns, (□) 20 ns, (▲) 30 ns.

is transferred along the side chain to the backbone by spin-diffusion.

Effect of correlation time

The degree of saturation transfer depends both on the correlation time and on the duration of the saturation. As both homonuclear cross- and auto-relaxation rates for macromolecules are dominated by $J(\Delta\omega) \approx J(0)$ (Equations 1 and 2), the rate constant for magnetization transfer is proportional to the rotational correlation time. Therefore, for a given degree of transfer, the saturation time should be inversely proportional to the correlation time. Figure 4C shows the saturation transfer to the amide protons in the SRY-DNA complex on saturating the imino protons for 1 s. For the large NOEs, the amount of transfer increases 5–7 fold on increasing the correlation time from 5 to 30 ns. In most cases, the amount of transfer begins to level off by about 20 ns (for 1 s irradiation). This is a corollary to the time-course observed for the 434Cro-DNA complex for a fixed correlation time (Figure 3B). Increasing the saturation time beyond a certain level brings little additional intensity, but may decrease the selectivity as spin-diffusion becomes more pronounced.

Saturation transfer within the MetJ-DNA complex

MetJ is a homodimer in which the primary DNA binding surface is a short antiparallel beta sheet formed from strands on the two subunits (Somers and Phillips, 1992). The distance criterion identifies the interface as comprising residues 12,13, 21–25 (β-sheet), and 52–55 (helix 2). Saturation of the anomeric protons of the DNA gives rise to substantial NOEs to these same residues (Figure 5A). NOEs elsewhere are very small, indicating that the DNA binding surface can be mapped in a straightforward manner. As for the other protein–DNA complexes studied here, the degree of saturation transfer to the protein is smaller when irradiating the DNA imino protons than when irradiating H1′ (Figure 5B). However, essentially the same set of residues in the protein is excited in both cases. Residues in the beta sheet make contact exclusively in the major groove of the DNA, whereas the other two groups of residues (12, 13 and 52–55) contact both grooves. Imino proton saturation gives only weak effects at residues 12 and 13 because the contacts are with sugar protons in the minor groove. In the beta sheet, the amino acid side chains interact with amino protons of Ade and Cyt in the major groove, and these protons are in good dipolar contact with the imino protons.

Saturation transfer within the U1A-RNA complex

The protein–DNA complexes are examples of different interaction surfaces with largely B-form duplex

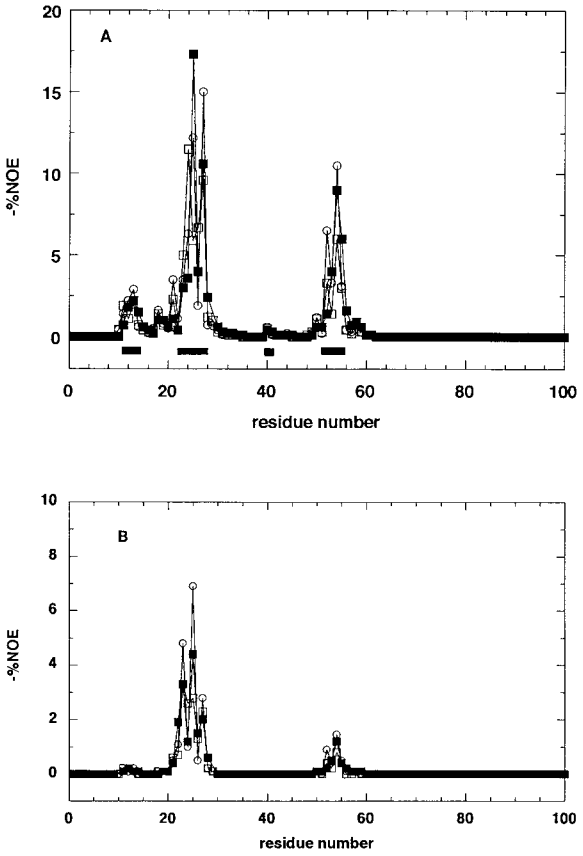


Figure 5. Cross-saturation in the MetJ-DNA complex. Simulations were carried out for a rigid body (except methyl groups) using an isotropic rotational correlation time of 20 ns. For clarity, only one subunit is shown. (A) NOEs for saturation of all H1′ for 1 s (♦) H_N, (□) Hα, (●) Hβ. (■) Residues with an atom within 3 Å of the DNA. (B) NOEs for saturation of all NH for 1 s (♦) H_N, (□) Hα, (○) Hβ.

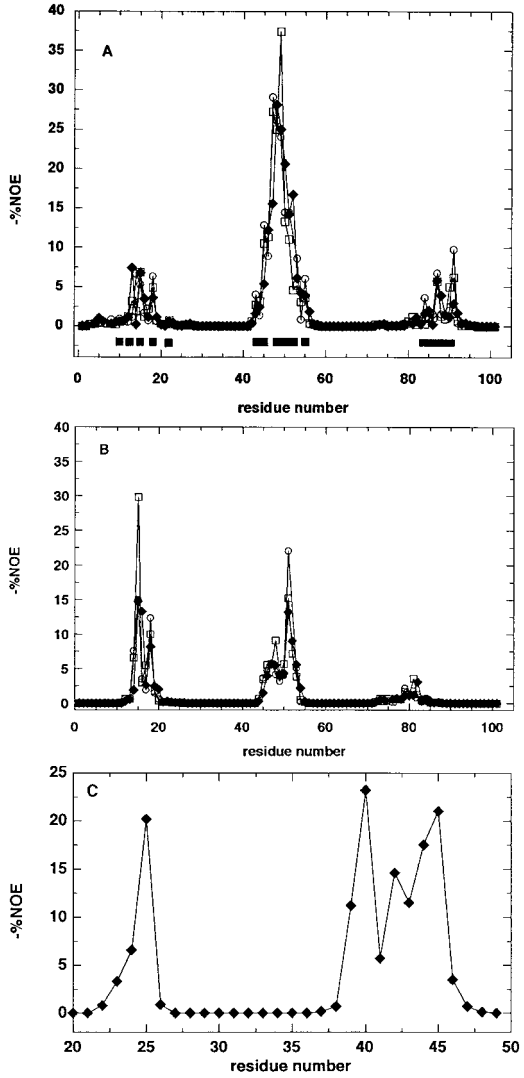


Figure 6. Cross-saturation in the U1A-RNA complex. Simulations were carried out for a rigid body (except methyl groups) using an isotropic rotational correlation time of 20 ns. (A) NOEs for 1 s saturation of H1′. (♦) H_N, (□) Hα, (○) Hβ. (■) Residues with an atom within 3 Å of the DNA. (B) NOEs for 1 s saturation of NH. (♦) H_N, (□) Hα, (○) Hβ. (C) NOEs to RNA H1′ resonances with 1 s saturation of the protein methyl protons.

DNA. In RNA–protein complexes, however, the RNA is usually a mixture of hairpin duplex (A-form) and exposed loop regions. Hence the nucleic acid geometry is very different from that found in DNA–protein complexes. According to the distance criterion, residues 12,15,18, 22, 43–55, 86–91 make close contact with the RNA.

Saturation of the anomeric protons (Figure 6A) or imino protons (Figure 6B) identifies the interface of the protein–RNA complex; the residues showing significant NOEs agree well with those found by the distance search, as shown by the filled squares in Figure 6A. However, the intensity pattern obtained with irradiation of the imino protons or the anomeric protons is significantly different. Thus, residues Tyr 12, Glu 18, Ser 27 and Arg 51 show substantial saturation transfer on irradiating the imino protons of the RNA, whereas on saturating the H1′ of the RNA, residues Ser 47 and Phe 55 show the greatest magnetization transfer, with smaller, but substantial transfer also to Tyr 12 and Glu 18, and to Ser 90. It is notable that for this complex, saturation of the imino protons gives larger NOEs to some residues of the protein than saturation of H1′. This clearly reflects the very different interaction mode. However, as most of the interactions are with an RNA loop, the imino protons may

exchange rapidly with water, an effect which was not taken into account in these simulations.

In contrast with DNA, RNA contains no methyl groups, and the sugar resonances are rarely found upfield of 3 ppm. Hence, it would be possible to carry out saturation transfer experiments in the reverse sense, starting with selective saturation of the upfield methyl and other resonances of the protein independently of the RNA, and monitoring the resulting transfer of saturation across the interface by recording spectra of the nucleic acid. Figure 6C shows the results of simulating the transfer of saturation from the protein methyl region to the RNA H1′, from which the RNA residues involved in the interface can be readily determined. Large NOEs are observed at residues G25, A39, U40, U41, G42, C43, A44 and C45, which are indeed the ones in closest contact with the protein (see above), and comprise the loop region connecting the two double-helical RNA loops. The observed effects originate with the methyl groups of just three residues, namely Leu 44, Leu 48 and Met 50.

Experimental analysis of saturation transfer within the Mbp1-DNA complex

Saturation of the imino protons of the DNA in the Mbp1-DNA complex gives rise to saturation transfer to the protein amides as shown in Figure 7. Numerous peaks are visible in the difference HSQC spectrum, and, as expected from the simulations above, are of comparatively low absolute intensity. The difference spectrum is very much less crowded than the control spectrum, indicating significant transfer only to a small subset of all backbone amides. The most intense peaks are actually found for residues in the wing region from 64 to 75. Other substantial transfer peaks are present for residues in the turn between the first two strand of the beta sheet (11–13), and some relatively weak NOEs to residues found in helix B. All of the identified cross-peaks in fact map on one face of the protein. Furthermore, cross-peaks from residues on the other side of the protein are systematically absent, e.g., Ile 82, Phe 82 and Gly 107. These cross-saturation peaks map well to the observed shift-difference profile (Taylor et al., 2000), and confirm the importance of the wing region for DNA binding. They also agree with the significant perturbations of shifts in the N-terminal region of the protein. Thus, the shifts and NH saturation transfer point to an extended DNA binding surface. Furthermore, a mutational analysis implicated lysine residues in the C-terminal region, and the shift profile suggested residues near 120–121 as being involved in DNA binding (Taylor et al., 2000). The cross-saturation experiment shows a weak cross peak to residue Lys121, and systematic absence of peaks for residues 100–110. These last residues are far from the face of the protein identified from the other cross-saturation peaks (Figure 7B). Thus the presence of contacts in the wing (Gln 74) and tail (Lys 121) is consistent with the mutational data.

The presence of antiparallel beta sheet in Mbp1 causes some (four) CαH to overlap the H1′ resonances. To avoid partial saturation of the protein, we have used selective excitation at a number of frequencies in the 5.5–6 ppm region. Stepping through the frequencies results in different sets of peaks in the difference HSQC spectrum. However, in general these correspond to subsets of peaks observed in the iminoproton experiment, indicating that the same set of residues is being excited. Unfortunately, full assignment of the DNA in the complex has not yet been achieved, so a more detailed mapping of the protein-DNA interactions cannot be made.

Discussion

The cross-saturation experiment is a valuable complement to shift mapping and hydrogen exchange experiments for determining the binding surface in protein-nucleic acid complexes. It requires only ^{15}N-labelled protein and is simple to carry out. Several points emerge from the simulations which are pertinent to the interpretation of such experiments.

The simulated NOE profiles show an excellent correspondence with the interface defined by the distance criterion. The good accuracy of the cross-saturation technique can be attributed to the comparatively slow rate of spin diffusion from a localized source (cf. Figures 1 and 3). Saturation of the imino protons is straightforward as they are generally well resolved from any protein peaks (typically 11–15 ppm). The resulting cross-saturation peaks are generally lower than those obtained from saturating H1′. Saturating H1′ may also excite downfield shifted CαH in the protein. These are often found in proteins containing antiparallel beta sheet structures but rare in DNA-binding proteins that have a high α-helical content. A search of the BioMagResBank database revealed that saturation of the region 5.4 ppm to 6.3 ppm would excite $> 90\%$ of the H1′ resonances, together with the (unwanted) perturbation of ca. 2 protein resonances per protein-

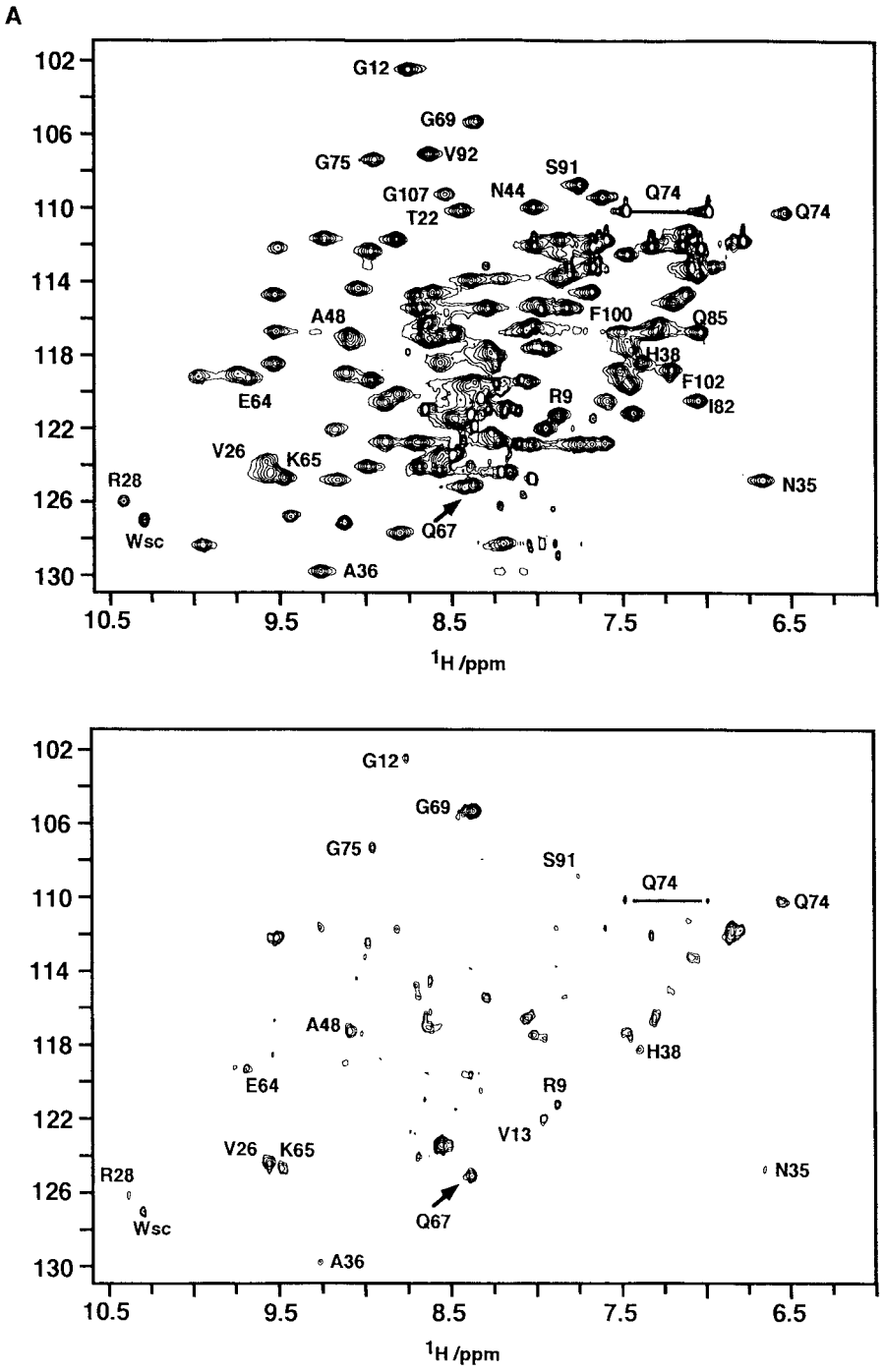
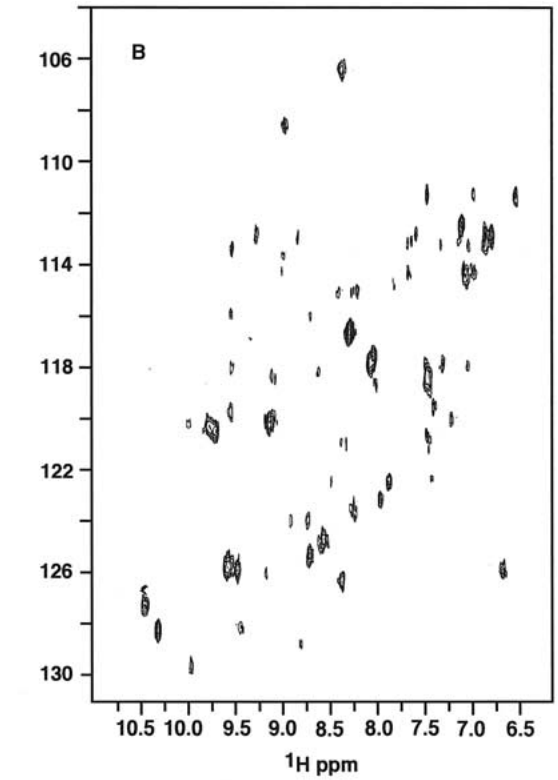


Figure 7. Cross-saturation in the Mbp1-DNA complex. Spectra were recorded at 14.1 T at 15 °C on a sample containing 0.6 mM Mbp1 and 0.65 mM duplex DNA. (A) saturation of NH for 2 s; upper control spectrum, lower difference spectrum. Particular assigned peaks discussed in the text are labelled. (B) CW excitation of H1′ at 6.23 ppm for 1 s. (C) Mbp1 structure with mapped NOEs and chemical shift differences.



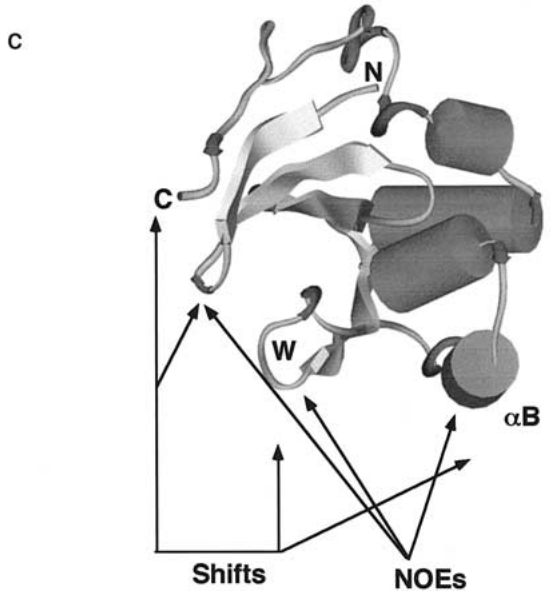


Figure 7. Continued.

nucleic acid complex on average. Three of the 11 assigned complexes had no protein resonances in this region. As the resonance positions are usually known in advance, the excitation profile can be tailored to avoid saturating them. Furthermore, excitation of one or two isolated protein protons is not too serious as the results can be interpreted in conjunction with the control experiment carried out on the free protein. If protein excitation remains problematic, band selective excitation would not be applicable, but useful information could still be obtained by selective excitation of individual anomeric protons. This gives the added benefit of determining which nucleotide is close to a particular amino acid residue in the protein. Excitation of H1′ also gives different relative intensities when compared with imino protons or the H5 of Cyt or Uri, reflecting their different spatial locations within the nucleic acid molecule. Combined with shift perturbation data, the cross-saturation experiments can provide detailed information about the protein-nucleic interface, and form the basis of docking calculations.

Our simulations suggest that one second of irradiation is a good compromise between sensitivity and selectivity for correlation times of 15–20 ns (cf. Figure 4C), and therefore for more slowly tumbling complexes, the saturation time should be decreased in proportion (or smaller complex, increased in proportion). For saturation of the imino protons, where spin-diffusion is slower, the saturation time can be increased 50–100% compared with the H1′ saturation experiment.

The simulations were carried out in the absence of water molecules. In practice, complexes in solution are extensively hydrated, but interfacial water can be expected to be less abundant than elsewhere owing to the tight juxtaposition of the interacting components. Relaxation by interfacial water requires close proximity of water protons to solute hydrogen atoms, and a residence time of the order of the rotational correlation time or longer. Cross-relaxation transfer between the nucleic acid and protein via bridging water molecules as observed in the trp repressor-operator complex and in a homeodomain–DNA complex seem unlikely because of their short residence time (Gruschus and Ferretti, 2001). The interface identified by our method is essentially one that involves direct protein–nucleic acid contacts. A second issue concerning the presence of water molecules is the possible leakage relaxation that would occur if a long lived water molecule was in close proximity to one of the interfacial protons involved in the magnetization transfer. However, in

the recent study of a homeodomain–DNA complex (Gruschus and Ferretti, 2001) such interfacial water was shown to give residence times in the range 0.4 to 3 ns, giving rise to only a small increase in local leakage relaxation. Simulations indicate that this has only a small effect on the observed transfer of saturation to the target protons. Hence, the ability of the method to identify the general nature of the interface should therefore not be impaired.

As the systems become larger, a major limitation becomes one of sensitivity owing to excessive line broadening. However, as ^{15}N-HSQC is the main detection method, this becomes a severe problem only for very large complexes. We have shown that good results can be obtained for complexes with a rotational correlation time of 22 ns (Mbp1-DNA at 15 °C, M = 23 kDa). Incorporation of TROSY-HSQC (Pervushin et al., 1997) is likely to be effective for complexes with rotational correlation times of up to ca. 50 ns. The increased spectral complexity implied by larger systems may require use mainly of the imino saturation to ensure good selectivity of excitation of the nucleic acid component.

Acknowledgements

This work was supported by the Medical Research Council of the U.K. NMR spectra were recorded at the MRC Biomedical NMR Centre, Mill Hill.

References

Albright, R.A. and Matthews, B.W. (1998) *J. Mol. Biol.*, **280**, 137–151.

Allain, F.H., Howe, P.W., Neuhaus, D. and Varani, G. (1997) *EMBO J.*, **16**, 5764–5772.

Beckmann, P., Martin, S.R. and Lane, A.N. (1993) *Eur. Biophys. J.*, **21**, 417–424.

Foster, M.P., Wuttke, D.S., Clemens, K.R., Jahnke, W., Radhakrishnan, I., Tennant, L., Raymond, M., Chung, J. and Wright, P.E. (1998) *J. Biomol. NMR*, **12**, 51–71.

Gruschus, J.M. and Ferretti, J. (2001) *J. Biomol. NMR*, **20**, 111–126.

Lane, A.N. (1988) *J. Magn. Reson.*, **78**, 425–439.

Lane, A.N. (1990) *Biochim. Biophys. Acta*, **1049**, 189–204.

Lane, A.N. and Fulcher, T. (1995) *J. Magn. Reson.*, **B107**, 34–42.

Lowndes, N.F., Johnson, A.L., Breeden, L. and Johnston, L.H. (1992) *Nature*, **357**, 505–508.

McIntosh, E.M., Atkinson, T., Storms, R.K. and Smith, M. (1991) *Mol. Cell Biol.*, **11**, 329–337.

McIntosh, P.B., Taylor, I.A., Frenkiel, T.A., Smerdon, S.J. and Lane, A.N. (2000) *J. Biomol. NMR*, **16**, 183–196.

Mondragon, A. and Harrison, S.C. (1991) *J. Mol. Biol.*, **219**, 321–334.

Paterson, Y., Englander, S.W. and Roder, H. (1990) *Science*, **249**, 755–759.

Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.

Ramos, A., Kelly, G., Hollingworth, D., Pastore, A. and Frenkiel, T. (2000) *J. Am. Chem. Soc.*, **122**, 11311–11314.

Somers, W.S. and Phillips, S.E. (1992) *Nature*, **359**, 387–393.

Takahashi, H., Nakanishi, T., Kami, K., Arata, Y. and Shimada, I. (2000) *Nat. Struct. Biol*, **7**, 220–223.

Taylor, I.A., McIntosh, P.B. Pala, P., Treiber, M.K., Howell, S., Lane, A.N. and Smerdon, S.J. (2000) *Biochemistry*, **39**, 3943–3954.

Taylor, I.A., Treiber, M.K., Olivi, L. and Smerdon, S.J. (1997) *J. Mol. Biol.*, **272**, 1–8.

Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.

Werner, M.H., Huth, J.R., Gronenborn, A.M. and Clore, G.M. (1995) *Cell*, **81**, 705–714.

Williamson, R.A., Carr, M.D., Frenkiel, T.A., Feeney, J. and Freedman, R.B. (1997) *Biochemistry*, **36**, 13882–13889.

Woessner, D.E. (1962) *J. Chem. Phys.*, **37**, 647–654.

Xu, R.M., Koch, C., Liu, Y., Horton, J.R., Knapp, D., Nasmyth, K. and Cheng, X. (1997) *Structure*, **5**, 349–358.