Structure of the chromatin binding (chromo) domain from mouse modifier protein 1

Linda J.Ball, Natalia V.Murzina,
R.William Broadhurst, Andrew R.C.Raine,
Sharon J.Archer¹, Francesca J.Stott,
Alexey G.Murzin², Prim B.Singh³,
Peter J.Domaille¹ and Ernest D.Laue⁴

Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK, ¹Du Pont Merck Pharmaceutical Company, Box 80328, Wilmington, DE 19880-0328, USA, ²Centre for Protein Engineering, MRC Centre, Hills Road, Cambridge CB2 2QH and ³The Babraham Institute, Babraham Hall, Babraham, Cambridge CB2 4AT, UK

⁴Corresponding author

The structure of a chromatin binding domain from mouse chromatin modifier protein 1 (MoMOD1) was determined using nuclear magnetic resonance (NMR) spectroscopy. The protein consists of an N-terminal three-stranded anti-parallel β -sheet which folds against a C-terminal α -helix. The structure reveals an unexpected homology to two archaeobacterial DNA binding proteins which are also involved in chromatin structure. Structural comparisons suggest that chromo domains, of which more than 40 are now known, act as protein interaction motifs and that the MoMOD1 protein acts as an adaptor mediating interactions between different proteins.

Keywords: chromo domain/homology/NMR spectroscopy/protein interaction motif/structure

Introduction

The chromatin organization modifier (chromo) domain has been defined as a 30–70 residue protein module found in a large number of proteins. So far more than 40 examples of the chromo domain are known (Aasland and Stewart, 1995; Koonin *et al.*, 1995) and the list is growing rapidly. Chromo domains were first identified in two *Drosophila* proteins, heterochromatin associated protein 1 (HP1) and Polycomb (Pc) (Paro and Hogness, 1991). HP1 has been shown to be involved in repression of gene expression in heterochromatin (Eissenberg *et al.*, 1990, 1992) and Pc is a member of the Polycomb group of transcriptional repressors which suppress expression of homeotic genes (Paro, 1990, 1993). Studies of both Pc (Messmer *et al.*, 1992) and HP1 (Platero *et al.*, 1995) have shown that chromo domains are responsible for targeting their respective proteins to their sites of action on chromatin. Similarly, the chromo domain in human retinoblastoma binding protein (RBP1) may deliver retinoblastoma protein (Rb), a global transcription factor, to particular sites (Szekely *et al.*, 1991; Koonin *et al.*, 1995). In addition, *Drosophila* male-specific lethal (MSL-3)

protein, an activator, might also be localized to chromatin via its chromo domain (Koonin *et al.*, 1995).

It is now apparent that the chromo domain family of proteins comprises at least two groups. The HP1-like proteins contain two conserved homologous domains (Clark and Elgin, 1992): an N-terminal chromo domain and a C-terminal 'shadow' chromo domain (Epstein *et al.*, 1992; Aasland and Stewart, 1995). Mutational analysis has shown that the N-terminal chromo domain of HP1 is sufficient to direct heterochromatin binding of a β -galactosidase fusion protein *in vivo*, whereas the C-terminal shadow domain is necessary for nuclear localization and heterochromatin binding (Powers and Eissenberg, 1993; Platero *et al.*, 1995). The two chromo domains of the HP1-like proteins may thus serve as adaptors bringing together different proteins in a multi-protein complex (Aasland and Stewart, 1995). In other cases, such as Pc, CHD-1 (Delmas *et al.*, 1993) and SWI6 (Lorentz *et al.*, 1994; Ekwall *et al.*, 1995), the chromo domain is a component of a much larger protein and is typically fused to other protein modules (Aasland and Stewart, 1995).

The HP1 gene was identified as a suppressor of position-effect variegation (PEV) (Eissenberg *et al.*, 1990). PEV occurs when a gene normally residing in euchromatin is placed next to a heterochromatic boundary after a chromosomal rearrangement. The gene becomes specifically silenced in some cells, but not in others, leading to a variegated phenotype. Suppressors or enhancers of PEV are likely to encode proteins that are involved in the formation or modification of chromatin structure (for a review see Reuter and Spierer, 1992). PEV is similar in many ways to other mechanisms of gene silencing, such as those involving the homeotic genes (Paro, 1993), mating type control in yeast (Laurenson and Rine, 1992), X-chromosome inactivation (Gartler and Riggs, 1983) and imprinting in mammalian cells (Solter, 1988).

In this paper we describe experiments that determined the domain structure of HP1-like heterochromatin-associated mouse chromatin modifier protein 1 (MoMOD1) (Singh *et al.*, 1991). The N-terminal chromo domain (MoMOD1-N, see Figure 1) was expressed in *Escherichia coli*, purified to homogeneity and its solution structure determined using nuclear magnetic resonance (NMR) spectroscopy. Unexpectedly, the structure is similar to that of the Sac7d and Sso7d DNA binding proteins from the archaeobacteria *Sulfolobus acidocaldarius* and *S.solfataricus* (Baumann *et al.*, 1994; Edmondson *et al.*, 1995). The structural comparisons suggest how the chromo domain may function as a protein interaction motif.

Results

Expression of the intact MoMOD1 protein

The full-length MoMOD1 protein (residues 1–185) was cloned into pET16b and expressed in *E.coli* BL21(DE3)

Fig. 1. Amino acid sequence of the MoMOD1 chromo domain showing its homology with other representative chromo domains. Selected conserved residues are coloured yellow (core hydrophobic), green (Gly and Pro) and blue (basic). See text for details.

Fig. 2. Limited proteolysis of MoMOD1: 16% tricine-SDS-PAGE separation of a 0.5% trypsin or 1% subtilisin digest of the MoMOD1 protein. (a) Lane 1, intact protein; lane 2, after 5 min incubation; lane 3, after 15 min incubation; lane 4, after 30 min incubation; lane 5, after 60 min incubation with subtilisin. (b) Lane 1, intact protein; lane 2, after 5 min incubation; lane 3, after 30 min incubation; lane 4, after 60 min incubation with trypsin. On the right of these panels, the results of the microsequencing and mass spectrometric analysis are summarized. The numbers represent the N- and C-terminal amino acids of the major components identified from the bands. (c) Amino acid sequence of the MoMOD1 protein showing the protease cut sites. The sequences of the protease-resistant N- and C-terminal domains are underlined. Up arrow, trypsin cleavage site; down arrow, subtilisin cleavage site.

as a histidine-tagged fusion protein. The protein was purified using nickel affinity chromatography, where it co-purified with small amounts of two shorter fragments of the protein. These may result from truncated transcription or translation or, perhaps, degradation by endogenous proteases. The shorter fragments did not interfere with interpretation of the limited proteolysis experiments presented below.

Limited proteolysis of the intact MoMOD1 protein

Before the proteolysis experiments were carried out, the stability of the protein was assessed under the conditions of the digest. This showed that the protein was not degraded. Digestion conditions were then varied over a wide range to determine which regions of the molecule are most susceptible to proteolysis. The proteolytic fragments were then separated using either tricine-SDS-PAGE (Figure 2a and b) or reversed phase HPLC (not shown)

and were identified using N-terminal amino acid sequencing and electrospray ionization mass spectroscopy. The N-terminus of each fragment was determined directly by N-terminal sequencing, whereas the C-terminus was identified indirectly using mass spectroscopy and the statistical program WEIGHT (W.Boucher, unpublished data); this program assigns a series of possible protein sequences to each mass spectral fragment. A summary of the results is presented in Figure 2c.

Expression of the MoMOD1-N chromo domain fragment

Based on the sequence alignment (Figure 1) and limited proteolysis data (Figure 2), a fragment containing residues 10–80 of MoMOD1 was expressed in *E.coli* in milligram quantities. The N-terminal sequence was HMVEEVL . . . , which is as expected with the addition of a histidine and methionine residue from the expression vector. Amino

Fig. 3. Location of the secondary structure in MoMOD1-N. Below the classification of the elements of secondary structure, the CSI line indicates the consensus chemical shift index for $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ nuclei. The stars below the amino acid sequence represent the location of slowly exchanging amide protons. In the next rows, filled and empty circles represent residues with $^3J_{\alpha\text{N}} > 9$ Hz and < 4 Hz respectively and squares represent residues for which χ_1 has been determined. Following this, three rows of solid bars represent the observed sequential $d_{\alpha\text{N}}$, $d_{\beta\text{N}}$ and d_{NN} NOE connectivities; the thickness of the bars indicates the intensities of the crosspeaks in the NOESY spectra. Further bars represent $d_{\alpha\text{N}}(i, i + 2)$, $d_{\text{NN}}(i, i + 2)$, $d_{\alpha\text{N}}(i, i + 3)$, $d_{\alpha\beta}(i, i + 3)$, $d_{\text{NN}}(i, i + 3)$ and $d_{\alpha\text{N}}(i, i + 4)$ NOE connectivities between the residues shown.

acid analysis showed the expected amino acid composition and electrospray ionization mass spectra identified one major species with a molecular mass of 8684.01; the calculated molecular mass is 8683.54. NMR studies of ^{15}N relaxation in the backbone amides (data not shown) show that the structured part of the protein has an overall rotational correlation time of 6.8 ns. In conjunction with the results of sedimentation analysis (data not shown), this is consistent with MoMOD1-N being monomeric in solution.

NMR assignment strategy and location of the secondary structure elements

The triple resonance assignment strategy employed for MoMOD1-N is similar to that outlined in Clowes *et al.* (1995). Assignments of the chemical shifts of ^1H , ^{13}C and ^{15}N nuclei in the protein backbone were made using a combination of three-dimensional CBCA(CO)NNH and CBCANNH experiments (Grzesiek and Bax, 1992a,b) and confirmed using three-dimensional ^{15}N -separated TOCSY-HMQC, NOESY-HMQC (Marion *et al.*, 1989a,b) and HNHB (Archer *et al.*, 1991) spectra. Aliphatic side chain ^1H and ^{13}C assignments were identified from four-dimensional HCC(CO)NNH (Clowes *et al.*, 1993) and three-dimensional ^{13}C -separated HCCH-TOCSY (Bax *et al.*, 1990a) experiments. Aromatic side chain signals were assigned by reference to two-dimensional CBHD, CBHE (Yamazaki *et al.*, 1993), DQF-COSY (Rance *et al.*, 1983) and NOESY (Kumar *et al.*, 1980) spectra.

Through-space nuclear Overhauser effect (NOE) connectivities from backbone amide to side chain protons were identified from a four-dimensional $^{13}\text{C}/^{15}\text{N}$ -separated NOESY experiment (Kay *et al.*, 1990). Amide–amide and side chain–side chain connections were obtained from three-dimensional ^{15}N - and ^{13}C -separated NOESY spectra

(Marion *et al.*, 1989b; Ikura *et al.*, 1990; Zuiderweg *et al.*, 1990). Vicinal coupling constants between amide and $^1\text{H}_\alpha$ spins, $^3J_{\alpha\text{N}}$, were determined by least squares fitting to the intensities of crosspeaks from a set of two-dimensional J-modulated ^{15}N - ^1H -COSY experiments (Billeter *et al.*, 1992), leading to estimates of the backbone angle ϕ by reference to a Karplus curve (Pardi *et al.*, 1984). Slowly exchanging amide protons were identified by recording ^1H - ^{15}N HSQC (Bax *et al.*, 1990b; Norwood *et al.*, 1990) spectra on a sample immediately after the H_2O solvent had been exchanged for D_2O .

The elements of secondary structure were then identified from $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts (Wishart and Sykes, 1994), patterns of short and intermediate range inter-residue NOEs, $^3J_{\alpha\text{N}}$ couplings and amide exchange rates (Wuthrich, 1986), as illustrated by the connectivity summary in Figure 3. The presence of strong $d_{\alpha\text{N}}$ NOEs, retarded amide exchange rates, positive values for the consensus chemical shift index (CSI) and large values of $^3J_{\alpha\text{N}}$ indicate that there are β -strands from residues 23 to 31, 36 to 42 and 51 to 54. Between residues 60 and 71 the $d_{\alpha\text{N}}(i, i + 3)$ and $d_{\alpha\beta}(i, i + 3)$ NOEs, the strong d_{NN} and weak $d_{\alpha\text{N}}$ NOEs, the slow amide exchange, the negative values for the consensus CSI and the small values of $^3J_{\alpha\text{N}}$ reveal a C-terminal α -helix.

Determination of the tertiary structure

In addition to the 329 sequential and 218 medium range inter-residue NOEs that were used to locate the elements of secondary structure in the previous section, 910 intra-residue and 329 long-range NOEs were identified. For certain side chains it was possible to identify the most populated χ_1 rotamer and make stereospecific assignments of $^1\text{H}_\beta$ or γ -methyl resonances. $^3J_{\text{C}'\text{C}}$ and $^3J_{\text{C}'\text{N}}$ couplings were determined for six Val, Ile and Thr residues pos-

sessing resolved γ -methyl signals from ^{15}N and ^{13}CO spin-echo difference experiments (Grzesiek *et al.*, 1993; Vuister *et al.*, 1993). Stereospecific $^1\text{H}_\beta$ assignments for a further nine residues could be made by analysis of peak intensities in three-dimensional ^{15}N -separated HMQC-TOCSY, HMQC-NOESY and HNHB spectra (Powers *et al.*, 1993).

Three dimensional structures of the chromo domain were produced essentially as described before (Kraulis *et al.*, 1994). Initial structures were computed using simulated annealing calculations from 1207 ^1H - ^1H distance restraints that could be determined unambiguously. These computations used the program X-PLOR (Brunger, 1992), starting from random globular initial structures (Nilges *et al.*, 1988, 1991). This set of preliminary structures was used to assist in assignment of the remaining NOE crosspeaks for which the assignments were ambiguous and in selection of acceptor carbonyl oxygen atoms for hydrogen bonds to the slowly exchanging backbone amide protons. The final round of calculations used a restraint list comprising 1937 distance restraints, of which 1786 were unambiguous and 151 semi-ambiguous (Nilges, 1995), 22 ϕ and 15 χ_1 dihedral angle restraints and 17 hydrogen bonds; the selected ensemble contained all of the 27 converged structures out of 40 calculated. The structures are well defined (see Figure 4A) except for the first 11 residues at the N-terminus and the last nine residues at the C-terminus, which are genuinely flexible in solution, as judged by NMR studies of ^{15}N relaxation in the backbone amides (data not shown). The chromo domain therefore corresponds to residues 21–71 of MoMOD1. The average r.m.s. differences from the mean structure are 0.66 Å ($\pm$ 0.14) for the backbone atoms and 1.14 Å ($\pm$ 0.17) for all the non-hydrogen atoms in the well-defined core of the molecule. If the less well-ordered loop region (residues 43–47) is included, these values increase to 0.96 Å ($\pm$ 0.27) and 1.52 Å ($\pm$ 0.29) respectively. No distance restraint was violated by >0.5 Å and no dihedral angle restraint by $>5.0^\circ$. The structures have good covalent geometry as judged by PROCHECK (Laskowski *et al.*, 1993) and non-bonded contacts (see Table I).

Analysis of the tertiary structure

The structure of the chromo domain is shown in Figures 4 and 5. The protein consists of an N-terminal three-stranded anti-parallel β -sheet, which packs against the C-terminal α -helix (Figure 4A). A hydrophobic core is formed by residues 23, 26, 38, 40, 42, 58, 60, 63, 64 and 67 (see Figure 4B), all of which are highly conserved in different chromo domains (Figure 1). Several of these residues (23, 40, 42, 58, 60 and 63) also lie at the bottom of a hydrophobic groove (see below) on the surface of the β -sheet below the α -helix. The first strand of the β -sheet is rather irregular and contains two bulges at residues 24 and 27 (Figure 6). The insertions and deletions in the sequences of the different chromo domains are all found in the loops. Insertions and deletions are found in the first and at the end of the second loop in the β -sheet (residues 33–34 and 47–48). Insertions are also found in the loop connecting the β -sheet to the α -helix (residues 59–60) (Aasland and Stewart, 1995). It is therefore likely that all the chromo domains will have similar three-dimensional structures.

Discussion

Limited proteolysis of MoMOD1 revealed two regions that are resistant to digestion; these correspond to the conserved sequences in the HP1-like proteins (Clark and Elgin, 1992). The results suggest that the MoMOD1 protein consists of two structural domains, residues 10–78 and 104–171, connected by an exposed linker. Our NMR studies of the MoMOD1-N chromo domain show that the structured region lies between residues 21 and 71, although it is possible that, in the intact MoMOD1 protein, residues 10–20 and 71–78 are protected.

The structure of the MoMOD1-N chromo domain explains the results of mutations in the chromo domains of both HP1 (Platero *et al.*, 1995) and Pc (Messmer *et al.*, 1992), which have been shown to disrupt gene silencing and/or chromatin binding. Three of these mutations, V23M in HP1, I26F and removal of residues 64/65 in Pc (all MoMOD1 numbering), are to residues in the hydrophobic core and they may therefore disrupt the structure of the chromo domain (Figure 4a). One mutation, Y21F in HP1, should not disrupt the structure, but this residue lies to one side of the hydrophobic groove, which we suggest mediates protein–protein interactions (see below). This tyrosine may therefore be important for the function of at least the HP1-like chromo domains. We note that the V23M mutation may additionally disrupt function, because this residue is also exposed in the hydrophobic groove.

Comparisons with other protein structures from the Brookhaven Protein Databank were carried out by eye and using the computer program DALI (Holm and Sander, 1993). Homologies to two different classes of protein were found. The first group includes the α and β chemokines, e.g. IL-8, MGSA/GRO and MIP-1 β , RANTES and MCP1, all of which are involved in protein–protein interactions with receptors (Leonard and Yoshimura, 1990). Most interestingly, the second group comprises small histone-like DNA binding proteins found in the archaeobacteria *S.acidocaldarius* and *S.solfataricus*. These proteins are also involved in the formation of chromatin structure (Dijk and Reinhardt, 1986) and they are more similar to the MoMOD1-N chromo domain. The r.m.s. deviations for the 31 C_α atoms, which comprise all of the conserved secondary structure, are 1.44 and 2.01 Å, when the chromo domain is compared with the DNA binding domains of Sac7d and Sso7d respectively (Baumann *et al.*, 1994; Edmondson *et al.*, 1995). For the DNA binding domains of Sac7d and Sso7d themselves, which are 70% identical, the r.m.s. deviation is 1.72 Å over the same region. (By comparison, the C_α r.m.s. deviations over the same region are 2.25 and 1.96 Å when the chromo domain is compared with the two chemokines IL-8 and MCP1 respectively.) In Figure 5a, a comparison of the structure of Sac7d with that of the chromo domain is shown. The structures are very similar, but the chromo domain lacks the N-terminal hairpin which binds in a hydrophobic groove below the α -helix in Sac7d. We believe that there may be a distant evolutionary relationship between the chromo domain and Sac7d, for the following reasons. First, in addition to conservation of the three-dimensional folds, there are detailed structural similarities as well. For example, the first bulge at residue 24 in the chromo domain is conserved in Sac7d, to facilitate

Fig. 4. (a) Stereoviews of the backbone (N, C α and C') of the 27 superimposed structures of MoMOD1-N (residues 20–73) that converged out of 40 computed from the NMR data. (b) Structure closest to the mean, showing the side chains of the conserved hydrophobic residues in the core and hydrophobic groove of the protein. The side chains are colour coded according to their r.m.s. deviation from the mean structure of all the non-hydrogen atoms in the set of calculated structures: red < 0.75 Å; light blue < 1.5 Å.

wrapping of the β -strand around the α -helix (Figure 6). The interaction of a basic residue in the β -sheet (residue 29) with residues 64 and 67 in the α -helix is also conserved (Figure 6), as is the loop between the third β -strand and the α -helix (Figure 5a). Second, both the HP1-like chromo domains and the archaeobacterial DNA binding proteins have a cluster of negatively charged residues adjacent to the N-terminus of the structured part of the chromo domain. Finally, distinct similarities in the control of transcription in archaeobacteria and eukaryotic cells have been found (Baumann *et al.*, 1995b) and both proteins are involved in the formation of chromatin structure (Dijk and Reinhardt, 1986).

The surface charge of MoMOD1-N and Sac7d is compared in Figure 5b. It will be seen that the charge

distribution on the exterior of the three-stranded β -sheet of Sac7d, which has been shown to be involved in binding non-specifically to the major groove of DNA (Baumann *et al.*, 1995a), is not conserved in the chromo domain. There are two clusters of basic residues on the surface of the chromo domain (residues 25, 41 and 43 and residues 29, 30, 33 and 35) but overall the charge is negative, suggesting that this class of chromo domain would not bind to DNA or RNA by itself. In support of this, MoMOD1-N shows no appreciable affinity for DNA–cellulose (data not shown). It is also interesting to note that mutations in one of the clusters (R29QR30Q and K31Q, both MoMOD1 numbering) had no influence on the activity of HP1 (Platero *et al.*, 1995). Instead, we believe that the structural homologies suggest that the

chromo domain is a protein interaction motif and that the structure of Sac7d provides a model for how the chromo domain might bind other proteins. In the structure of the chromo domain there is a hydrophobic groove around the protein, where the N-terminal hairpin would bind in Sac7d, and we speculate that protein–protein interactions occur via this groove (Figure 5c). This hypothesis is supported by previous studies which suggest that mutation of residues at the edge of this groove (Y21F) and within the groove itself (V23M) disrupt gene silencing in HP1 (Platero *et al.*, 1995). It has been proposed that gene silencing might be controlled by tyrosine phosphorylation at this site (Platero *et al.*, 1995; Madireddi *et al.*, 1996) and phosphorylation should also affect protein–protein interactions in this groove. Strong support for the notion that the chromo domain mediates protein–protein interactions also comes from the results of two-hybrid screens, where both Le Douarin *et al.* (1996) and ourselves (N.Murzina, K.Johnson, A.De Smet and E.D.Laue, submitted for publication) have isolated different proteins that interact with the chromo domains of either the HP1 α or MoMOD1 proteins respectively.

It has been proposed that the C-terminal domain may be related to the N-terminal chromo domain in HP1-like proteins (Epstein *et al.*, 1992) and the term ‘shadow’ domain has been suggested (Aasland and Stewart, 1995). Sequence comparisons based on the structure of the MoMOD1-N chromo domain provide direct experimental

support for this hypothesis. If the N- and C-terminal domains have a similar fold we would expect that residues in the hydrophobic core and key residues required to make the turns in the β -sheet would be conserved between the N-terminal chromo and C-terminal shadow domains. Examination of Figure 1 shows this to be the case. The residues in the hydrophobic core and the glycine residues in loops 1 and 2 (Gly34 and Gly44) are indeed conserved between the N-terminal chromo and many of the

Table I. Structural statistics for the final ensemble of 27 refined structures of MoMOD1-N

Root mean square deviations from restraints and idealized geometry		
	$\langle SA \rangle^a$	$(SA)_c^b$
NOE distances (Å)	0.010 ± 0.003	0.009
Dihedral angles (°)	0.61 ± 0.29	0.53
Bonds (Å)	0.0016 ± 0.0003	0.0016
Angles (°)	0.38 ± 0.08	0.39
Impropers (°)	0.22 ± 0.04	0.22
Final energy		
E_{L-J} (kJ/mol)	-119.2 ± 26.5	-119.8
Assessment of backbone quality according to the Ramachandran plot		
	All structures	Closest to mean
Most favoured region	63.9%	66.7%
Additionally allowed region	28.7%	25.0%

^aRepresents the average r.m.s. deviations for the ensemble.

^bRepresents values for the final structure that is closest to the mean.

Fig. 5. Comparison of the structure of the MoMOD1-N chromo domain with that of the DNA binding domain of Sac7d (Edmondson *et al.*, 1995). (a) MOLSCRIPT (Kraulis, 1991) plots, in the same orientation as those shown in Figure 4, showing a comparison, of the structures closest to the mean, of the chromo domain from MoMOD1 (left) with that of the DNA binding domain from Sac7d (right); the side chains of residues where mutation affects gene silencing and/or chromatin binding (Messmer *et al.*, 1992; Platero *et al.*, 1995) are shown. (b) Electrostatic potential at the protein surfaces contoured and colour coded at -4.8 kT (red) and $+4.8$ kT (blue). The potential was calculated and displayed with the program GRASP (Gilson *et al.*, 1988; Nicholls *et al.*, 1991). Compared with the plots in Figures 4 and 5a and c, the structures in (b) have been rotated by $\sim 130^\circ$ so that the β -sheet is now facing out of the page. (c) A similar surface, but now showing the putative peptide binding site in the chromo domain and binding of the N-terminal hairpin in Sac7d.

Fig. 6. A comparison of the secondary structure of the chromo domain (left) with that of the DNA binding domain of Sac7d (Edmonson *et al.*, 1995) (right). Dashed lines represent hydrogen bonds, thick lines conserved secondary structure and circles hydrophobic residues in the protein core. A conserved arginine that interacts with residues 64 and 67 in the α -helix (MoMOD1 numbering) is also shown.

C-terminal shadow domains. It therefore seems very likely that proteins like MoMOD1 do act as adaptors mediating interactions between different proteins (Aasland and Stewart, 1995; Koonin *et al.*, 1995). However, some structural differences are likely, suggesting that the two domains will have distinct roles. For example, the conserved Pro 54, which moves the main chain out of the plane of the β -sheet, is typically replaced by lysine in the shadow domain. Moreover, many surface residues that are conserved in the chromo domain are different in the shadow domain, making that domain notably more hydrophobic. The results of two-hybrid screens by both Le Douarin *et al.* (1996) and ourselves (N.Murzina, K.Johnson, A.De Smet and E.D.Laue, submitted for publication) indicate that the N- and C-terminal domains in HP1-like proteins will have a distinct binding specificity. Most interestingly, and consistent with their possible role in recognizing and binding to different proteins, the size and shape of the residues lining the hydrophobic cleft (residues 23, 40, 42, 58, 60 and 63) are not conserved between the different classes of chromo domain (Figure 1).

The importance of chromo domains and proteins involved in the control of heterochromatin condensation has been re-affirmed by two sets of recent results. First, it has been shown that there is a molecular link between programmed DNA degradation and heterochromatin assembly (Madireddi *et al.*, 1996). Second, two groups have shown that heterochromatin contains pairing elements that act additively to initiate, or maintain, the proper alignment of achiasmatic chromosomes in meiosis I (Dernburg *et al.*, 1996; Karpen *et al.*, 1996). Site-directed mutagenesis, based on the NMR-derived solution structure reported here, will allow characterization of the residues that mediate the binding and biological activity of these important protein modules and should be of use in further

understanding the role of chromo domain proteins in formation of chromatin structure and control of transcription.

Materials and methods

Plasmids and bacterial strains

The MoMOD1 cDNA, cloned into the Bluescript SK+/- vector as an *EcoRI* fragment (Singh *et al.*, 1991), was used for construction of both pET-MoMOD1 and pET-MoMOD1-N (see below). Bacterial strains used were *E.coli* TG1 recO [genotype K12 $\Delta(lac-pro)$ *supE thi hsd D5/F' tra D36 proA⁺B⁺ lac^q Lac Z Δ M15 recO::Tn5*] (constructed by Dr P.Oliver, Department of Genetics, Cambridge, UK) and BL21(DE3) (Novagen).

Construction of pET-MoMOD1 and pET-MoMOD1-N

PCR was used to subclone the regions of the MoMOD1 cDNA corresponding to amino acids residues 1–185 (for MoMOD1) and 10–80 (for MoMOD1-N). Primers 1 (5'-GGATCCAT ATG GGG AAA AAG CA-3') and 2 (5'-GGATCCAT ATG GTG GAG GAG-3') introduced an *NdeI* site encompassing the ATG start codon at the 5'-end of the coding sequence; primers 3 (5'-GGATCCGAATTC TCA ATT CTT GTC-3') and 4 (5'-GGATCCGAATTC TCA TGA CTT ATC TG-3') introduced a stop codon and either an *EcoRI* or *BamHI* restriction site at the 3'-end of the fragment. The PCR products were digested with *NdeI* and either *EcoRI* or *BamHI* and cloned into pET16b (Novagen). The ligation products were transformed into TG1 recO and the DNA sequence of the insert verified.

Expression and purification of MoMOD1 and MoMOD1-N

Escherichia coli BL21(DE3) cells were transformed with either pET-MoMOD1 or pET-MoMOD1-N and grown at 37°C in 2 $\times$ TY medium containing 100 μ g/ml ampicillin. Expression was induced at an OD₆₀₀ of 0.70 with 1 mM isopropyl β -D-thiogalactopyranoside (IPTG) and the cells were incubated for a further 3 h. All the following procedures were carried out at 0–4°C and were monitored using 16% tricine-SDS-PAGE (Schagger and von Jagow, 1987). Cell pellets were resuspended in lysis buffer [0.5 M NaCl, 20 mM Tris, pH 7.9, and 1 mM phenylmethylsulfonyl fluoride (PMSF)] and lysed with a French press at 1000 p.s.i. The lysate was centrifuged at 35 000 r.p.m. for 1 h in a Beckman 42.1 Ti rotor and the supernatant was then loaded onto a 10 ml nickel-IDA column (Sigma Fast Flow) equilibrated with buffer A (100 mM NaCl, 20 mM Tris, pH 7.9, and 0.05% Na₃). The column was washed with buffer A and then with 60 mM imidazole in buffer A. The protein was eluted using a step gradient of imidazole (40 ml 60–500 mM, 30 ml 500 mM, 20 ml 500 mM–1 M) at a flow rate of 2 ml/min and monitored using UV spectroscopy at 254 nm. The protein eluted between 300 and 500 mM imidazole. The pooled peak fractions containing the protein were concentrated using a stirred cell (Amicon) with a YM3 ultrafiltration membrane. The buffer was exchanged for 100 mM NaCl, 50 mM Tris, pH 8.0, and 1 mM CaCl₂ using an Econo-Pack 10DG column (BioRad). The N-terminal histidine tag was removed using 1:100 (w/w) Factor Xa (Boehringer) in an overnight incubation at 22°C. The solution was then reloaded onto the nickel-IDA affinity column to remove the histidine tag and any remaining uncleaved protein.

Proteolytic digestion

Proteolytic reactions contained 50 μ g protein in 1 mM triethanolamine, pH 7.5. Either trypsin or subtilisin (Sigma) was added to a final concentration of either 20, 10, 5, 1 or 0.2% w/w. The reactions were carried out at 26°C and aliquots of the reaction were taken at 5, 30, 60 and 120 min and mixed with PMSF and tricine-SDS-PAGE loading buffer. Proteolytic reactions were then repeated using selected enzyme concentrations and incubated for a single chosen time point. The digestion products were separated using tricine-SDS-PAGE or reverse phase HPLC.

Electroblotting and N-terminal sequencing

Protein fragments were separated using tricine-SDS-PAGE and transferred by semi-dry electroblotting (Pharmacia LKB) to a Problott membrane. The running buffer included 2 mM mercaptoacetic acid and a Trizma transfer buffer of 25 mM Tris, 190 mM glycine, 10% methanol and 0.02% SDS. The membrane was stained with 0.1% Coomassie R250 in 50% methanol and 1% acetic acid for 5 min, destained with 50% methanol and washed in water. The regions of the membrane containing

the protein bands of interest were subjected to automated N-terminal amino acid microsequence analysis.

Reverse phase HPLC (RP HPLC) and mass spectrometry

RP HPLC of peptide samples for electrospray ionization mass spectrometry was performed using a Hewlett Packard 1090 liquid chromatograph. The peptide samples were purified using a GP18 Octadecyl (C18) column and eluted in a gradient of 0–100% acetonitrile containing 0.1% (v/v) trifluoroacetic acid.

NMR spectroscopy

Uniformly $^{13}\text{C}/^{15}\text{N}$ - and ^{15}N -labelled MoMOD1-N proteins were prepared for the NMR experiments by growing *E. coli* BL21(DE3) (Novagen) harbouring the expression plasmid in a MOPS minimal medium containing either 10% ^{15}N - or $^{13}\text{C}/^{15}\text{N}$ -labelled Celltone (Martek) and $[\text{N}^{15}]\text{H}_4\text{Cl}$, either with or without $[\text{C}^{13}]\text{glucose}$. The protein samples were purified as described above. Four protein samples of MoMOD1-N were prepared at pH 6.0 in a buffer containing 10 mM KH_2PO_4 , 25 mM NaCl and 0.05% NaN_3 : a 2.0 mM ^{15}N -labelled sample, a 2.7 mM $^{13}\text{C}/^{15}\text{N}$ -labelled sample, a 3.5 mM unlabelled sample, all in 90% H_2O , 10% D_2O , and a 1.0 mM $^{13}\text{C}/^{15}\text{N}$ -labelled sample in 100% D_2O .

All spectra were acquired using a Bruker AMX600 spectrometer at 298 K as described in either Kraulis *et al.* (1994) or the original references. For the identification of slowly exchanging amide protons, the ^{15}N -labelled 90% H_2O , 10% D_2O sample was buffer exchanged into 99.8% D_2O by 20-fold dilution followed by concentration using a stirred cell (Amicon) with a YM3 ultrafiltration membrane. All the data were processed using the AZARA software package (W.Boucher, unpublished data) and analysed with the program ANSIG 3.3 (Kraulis *et al.*, 1994) on Silicon Graphics Iris workstations. Both of these programs are available by anonymous ftp to ftp.bio.cam.ac.uk in the directory ~ftp/pub.

Structure calculations

The structures were calculated from the NMR data using the program X-PLOR (Brunger, 1992), but with minor modification to the files sa.inp and refine.inp from the X-PLOR 3.1 release (M.Nilges, personal communication). Crosspeaks which could not be assigned unambiguously were included after the manner of Nilges (1995), restraining the r^{-6} sum of the distances between contributing atoms. The structures and restraint files will be deposited in the Brookhaven Protein Databank.

Acknowledgements

Protein microsequencing, oligonucleotide synthesis, amino acid analysis and electrospray ionization mass spectrometry were carried out in the Protein and Nucleic Acid Chemistry Facility in the Department of Biochemistry, University of Cambridge. We thank Dan Camac (Du Pont Merck Pharmaceutical Company) for the sedimentation analysis, Len Packman for the N-terminal sequencing and mass spectrometry and Michael Nilges (EMBL) for protocols for the structure calculations. This work was supported by a grant from the Wellcome Trust. The Cambridge Centre for Molecular Recognition is supported by the BBSRC and the Wellcome Trust. A.G.M. is a MRC Senior Research Fellow.

References

Aasland, R. and Stewart, A.F. (1995) The chromo shadow domain, a second chromo domain in heterochromatin-binding protein-1, HP-1. *Nucleic Acids. Res.*, **23**, 3168–3173.

Archer, S.J., Ikura, M., Torchia, D.A. and Bax, A. (1991) An alternative 3D-NMR technique for correlating backbone ^{15}N with sidechain H_β -resonances in larger proteins. *J. Magn. Resonance*, **95**, 636–641.

Baumann, H., Knapp, S., Lundback, T., Ladenstein, R. and Hard, T. (1994) Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus solfataricus*. *Nature Struct. Biol.*, **1**, 808–819.

Baumann, H., Knapp, S., Karshikoff, A., Ladenstein, R. and Hard, T. (1995a) DNA-binding surface of the Sso7d protein from *Sulfolobus solfataricus*. *J. Mol. Biol.*, **247**, 840–846.

Baumann, P., Qureshi, S.A. and Jackson, S.P. (1995b) Transcription—new insights from studies on archaea. *Trends Genet.*, **11**, 279–283.

Bax, A., Clore, G.M. and Gronenborn, A.M. (1990a) ^1H - ^1H correlation via isotropic mixing of ^{13}C magnetization, a new 3-dimensional approach for assigning ^1H and ^{13}C spectra of ^{13}C enriched proteins. *J. Magn. Resonance*, **88**, 425–431.

Bax, A., Ikura, M., Kay, L.E., Torchia, D.A. and Tschudin, R. (1990b)

Comparison of different modes of 2-dimensional reverse-correlation NMR for the study of proteins. *J. Magn. Resonance*, **86**, 304–318.

Billeter, M., Neri, D., Otting, G., Qian, Y.Q. and Wuthrich, K. (1992) Precise vicinal coupling constants $^3J_{\text{NH}}$ in proteins from nonlinear fits of J-modulated ^{15}N - ^1H -COSY experiments. *J. Biomol. NMR*, **2**, 257–274.

Brunger, A. (1992) X-PLOR, Version 3.1, A System for X-Ray Crystallography and NMR. Yale University Press, New Haven, CT.

Clark, R.F. and Elgin, S.C.R. (1992) Heterochromatin protein-1, a known suppressor of position-effect variegation, is highly conserved in *Drosophila*. *Nucleic Acids Res.*, **20**, 6067–6074.

Clowes, R.T., Boucher, W., Hardman, C.H., Domaille, P.J. and Laue, E.D. (1993) A 4D HCC(CO)NNH experiment for the correlation of aliphatic sidechain and backbone resonances in $^{13}\text{C}/^{15}\text{N}$ labeled proteins. *J. Biomol. NMR*, **3**, 349–354.

Clowes, R.T., Crawford, A., Raine, A.R.C., Smith, B.O. and Laue, E.D. (1995) Improved methods for structural studies of proteins using NMR spectroscopy. *Curr. Opin. Biotechnol.*, **6**, 81–88.

Delmas, V., Stokes, D.G. and Perry, R.P. (1993) A mammalian DNA-binding protein that contains a chromo domain and an SNF2/SWI2-like helicase domain. *Proc. Natl Acad. Sci. USA*, **90**, 2414–2418.

Dernburg, A.F., Sedat, J.W. and Hawley, R.S. (1996) Direct evidence of a role for heterochromatin in meiotic chromosome segregation. *Cell*, **86**, 135–146.

Dijk, J. and Reinhardt, R. (1986) In Gualerzi, C. and Pon, C. (eds), *Bacterial Chromatin*. Springer Verlag, Berlin, Germany.

Edmondson, S.P., Qui, L. and Shriver, J.W. (1995) Solution structure of the DNA-binding protein Sac7d from the hyperthermophile *Sulfolobus acidocaldarius*. *Biochemistry*, **34**, 13289–13304.

Eisenberg, J.C., James, T.C., Foster-Hartnett, D.M., Hartnett, T., Ngan, V. and Elgin, S.C.R. (1990) Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA*, **87**, 9923–9927.

Eisenberg, J.C., Morris, G.D., Reuter, G. and Hartnett, T. (1992) The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics*, **131**, 345–352.

Ekwall, K., Javerzat, J.-P., Lorentz, A., Schmidt, H., Cranston, G. and Allshire, R.C. (1995) The chromodomain protein Swi6: a key component at fission yeast centromeres. *Science*, **269**, 1429–1431.

Epstein, H., James, T.C. and Singh, P.B. (1992) Cloning and expression of *Drosophila* HP1 homologs from a mealybug, *Planococcus citri*. *J. Cell Sci.*, **101**, 463–474.

Gartler, S.M. and Riggs, A.D. (1983) Mammalian X-chromosome inactivation. *Annu. Rev. Genet.*, **17**, 155–190.

Gilson, M.K., Sharp, K.A. and Honig, B. (1988) Calculation of the total electrostatic energy of a macromolecular system—solvation energies, binding energies and conformational analysis. *J. Comput. Chem.*, **9**, 327–335.

Grzesiek, S. and Bax, A. (1992a) Correlating backbone amide and sidechain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.*, **114**, 6291.

Grzesiek, S. and Bax, A. (1992b) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J. Magn. Resonance*, **99**, 201–207.

Grzesiek, S., Vuister, G.W. and Bax, A. (1993) A simple and sensitive experiment for measurement of J_{CC} couplings between backbone carbonyl and methyl carbons in isotopically enriched proteins. *J. Biomol. NMR*, **3**, 487–493.

Holm, L. and Sander, C. (1993) Protein-structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.

Ikura, M., Kay, L.E. and Bax, A. (1990) A novel approach for sequential assignment of ^1H , ^{13}C and ^{15}N spectra of larger proteins—heteronuclear triple-resonance 3-dimensional NMR-spectroscopy—application to calmodulin. *Biochemistry*, **29**, 4659–4667.

Karpen, G.H., Le, M.-H. and Le, H. (1996) Centric heterochromatin and the efficiency of achiasmate disjunction in *Drosophila* female meiosis. *Science*, **273**, 118–122.

Kay, L.E., Clore, G.M., Bax, A. and Gronenborn, A.M. (1990) 4-Dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1 β in solution. *Science*, **249**, 411–414.

Koonin, E.V., Zhou, S.B. and Lucchesi, J.C. (1995) The chromo superfamily—new members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin. *Nucleic Acids. Res.*, **23**, 4229–4233.

- Kraulis,P.J. (1991) Molscript—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.
- Kraulis,P.J., Domaille,P.J., Campbell-Burk,S.L., Van Aken,T. and Laue,E.D. (1994) Solution structure and dynamics of Ras p21-GDP determined by heteronuclear three- and four-dimensional NMR spectroscopy. *Biochemistry*, **33**, 3515–3531.
- Kumar,A., Ernst,R.R. and Wuthrich,K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.*, **95**, 1–6.
- Laskowski,R.A., Macarthur,M.W., Moss,D.S. and Thornton,J.M. (1993) Procheck—a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.
- Laurenson,P. and Rine,J. (1992) Silencers, silencing and heritable transcriptional states. *Microbiol. Rev.*, **56**, 543–560.
- Leonard,E.J. and Yoshimura,T. (1990) Human monocyte chemoattractant protein-1 (MCP-1). *Immunol. Today*, **11**, 97–101.
- Le Douarin,B., Nielsen,A.L., Garnier,J.-M., Ichinose,H., Jeanmougin,F., Losson,R. and Chambon,P. (1996) A possible involvement of TIF1 α and TIF1 β in the epigenetic control of transcription by nuclear receptors. *EMBO J.*, **15**, 6701–6715.
- Lorentz,A., Ostermann,K., Fleck,O. and Schmidt,H. (1994) Switching gene Swi6, involved in repression of silent mating-type loci in fission yeast, encodes a homologue of chromatin-associated proteins from *Drosophila* and mammals. *Gene*, **143**, 139–143.
- Madireddi,M.T., Coyne,R.S., Smothers,J.F., Mickey,K.M., Yao,M.-C. and Allis,C.D. (1996) PDD1P, a novel chromo domain-containing protein, links heterochromatin assembly and DNA elimination in tetrahymena. *Cell*, **87**, 75–84.
- Marion,D., Ikura,M., Tschudin,R. and Bax,A. (1989a) Rapid recording of 2D NMR-spectra without phase cycling—application to the study of hydrogen exchange in proteins. *J. Magn. Resonance*, **85**, 393–399.
- Marion,D., Driscoll,P.C., Kay,L.E., Wingfield,P.T., Bax,A., Gronenborn,A.M. and Clore,G.M. (1989b) Overcoming the overlap problem in the assignment of ^1H -NMR spectra of larger proteins by use of 3-dimensional heteronuclear ^1H - ^{15}N Hartmann–Hahn multiple quantum coherence and nuclear Overhauser multiple quantum coherence spectroscopy—application to interleukin-1 β . *Biochemistry*, **28**, 6150–6156.
- Messmer,S., Franke,A. and Paro,R. (1992) Analysis of the functional role of the Polycomb chromo domain in *Drosophila melanogaster*. *Genes Dev.*, **6**, 1241–1254.
- Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association—insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.
- Nilges,M. (1995) Calculation of protein structures with ambiguous distance restraints—automated assignment of ambiguous NOE crosspeaks and disulphide connectivities. *J. Mol. Biol.*, **245**, 645–660.
- Nilges,M., Clore,G.M. and Gronenborn,A.M. (1988) Determination of three-dimensional structures of protons from interproton distance data by dynamical simulated annealing from a random array of atoms. *FEBS Lett.*, **239**, 129–136.
- Nilges,M., Kuszewski,J. and Brünger,A.T. (1991) In Hoch,J.C. (ed.), *Computational Aspects of the Study of Biological Macromolecules by NMR*. Plenum Press, New York, NY.
- Norwood,T.J., Boyd,J., Heritage,J.E., Soffe,N. and Campbell,I.D. (1990) Comparison of techniques for ^1H -detected heteronuclear ^1H - ^{15}N spectroscopy. *J. Magn. Resonance*, **87**, 488–501.
- Pardi,A., Billeter,M. and Wuthrich,K. (1984) Calibration of the angular dependence of the amide proton-C α -proton coupling constants, $^3J_{\text{HN}\alpha}$, in a globular protein—use of $^3J_{\text{HN}\alpha}$ for identification of helical secondary structure. *J. Mol. Biol.*, **180**, 741–751.
- Paro,R. (1990) Imprinting a determined state into the chromatin of *Drosophila*. *Trends Genet.*, **6**, 416–421.
- Paro,R. (1993) Mechanisms of heritable gene repression during development in *Drosophila*. *Curr. Opin. Cell Biol.*, **5**, 999–1005.
- Paro,R. and Hogness,D. (1991) The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. *Proc. Natl Acad. Sci. USA*, **88**, 263–267.
- Platero,J.S., Hartnett,T. and Eissenberg,J.C. (1995) Functional analysis of the chromo domain of HP1. *EMBO J.*, **14**, 3977–3986.
- Powers,J.A. and Eissenberg,J.C. (1993) Overlapping domains of the heterochromatin-associated protein HP1 mediate nuclear localization and heterochromatin binding. *J. Cell Biol.*, **120**, 291–299.
- Powers,R., Garrett,D.S., March,C.J., Frieden,E.A., Gronenborn,A.M. and Clore,G.M. (1993) The high-resolution, 3-dimensional solution structure of human interleukin-4 determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry*, **32**, 6744–6762.
- Rance,M., Sorensen,O.W., Bodenhausen,G., Wagner,G., Ernst,R.R. and Wuthrich,K. (1983) Improved spectral resolution in COSY ^1H -NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.*, **117**, 479–485.
- Reuter,G. and Spierer,P. (1992) Position effect variegation and chromatin proteins. *BioEssays*, **14**, 605–612.
- Schagger,H. and von Jagow,G. (1987) Tricine sodium dodecyl sulfate polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 kDa to 100 kDa. *Anal. Biochem.*, **166**, 368–379.
- Singh,P.B., Miller,J.R., Pearce,J., Kothary,R., Burton,R.D., Paro,R., James,T.C. and Gaunt,S.J. (1991) A sequence motif found in a *Drosophila* heterochromatin protein is conserved in plants and animals. *Nucleic Acids. Res.*, **19**, 789–794.
- Solter,D. (1988) Differential imprinting and expression of maternal and paternal genomes. *Annu. Rev. Genet.*, **22**, 127–146.
- Szekely,L., Uzvolgyi,E., Jiang,W.-Q., Durko,M., Wiman,K.G., Klein,G. and Sumegi,J. (1991) Subcellular localization of the retinoblastoma protein. *Cell Growth Differentiation*, **2**, 287–295.
- Vuister,G.W., Wang,A.C. and Bax,A. (1993) Measurement of 3-bond nitrogen–carbon J-couplings in proteins uniformly enriched with ^{15}N and ^{13}C . *J. Am. Chem. Soc.*, **115**, 5334–5335.
- Wishart,D.S. and Sykes,B.D. (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol.*, **239**, 363–392.
- Wuthrich,K. (1986) *NMR of Proteins and Nucleic Acids*. Wiley, New York, NY.
- Yamazaki,T., Forman-Kay,J.D. and Kay,L.E. (1993) 2-Dimensional NMR experiments for correlating $^{13}\text{C}_\beta$ and $^1\text{H}_{\delta\epsilon}$ chemical shifts of aromatic residues in ^{13}C -labeled proteins via scalar couplings. *J. Am. Chem. Soc.*, **115**, 11054–11055.
- Zuiderweg,E.R.P., McIntosh,L.P., Dahlquist,F.W. and Fesik,S.W. (1990) 3-Dimensional ^{13}C -resolved proton NOE spectroscopy of uniformly ^{13}C -labeled proteins for the NMR assignment and structure determination of larger molecules. *J. Magn. Resonance*, **86**, 210–216.

Received on December 24, 1996; revised on January 27, 1997

Note added in proof

After this paper had been accepted we became aware of work that shows that HPI also interacts with the lamin B receptor [Ye,Q. and Worman,H.J. (1996) *J. Biol. Chem.*, **271**, 14653–14656].