Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis

Nathan P. Cowieson*†, Janet F. Partridge†‡, Robin C. Allshire‡ and Paul J. McLaughlin*

Background: Proteins such as HP1, found in fruit flies and mammals, and Swi6, its fission yeast homologue, carry a chromodomain (CD) and a chromo shadow domain (CSD). These proteins are required to form functional transcriptionally silent centromeric chromatin, and their mutation leads to chromosome segregation defects. CSDs have only been found in tandem in proteins containing the related CD. Most HP1-interacting proteins have been found to associate through the CSD and many of these ligands contain a conserved pentapeptide motif.

Results: The 1.9 Å crystal structure of the Swi6 CSD is presented here. This reveals a novel dimeric structure that is distinct from the previously reported monomeric nuclear magnetic resonance (NMR) structure of the CD from the mouse modifier 1 protein (MoMOD1, also known as HP1β or M31). A prominent pit with a non-polar base is generated at the dimer interface, and is commensurate with binding an extended pentapeptide motif. Sequence alignments based on this structure highlight differences between CDs and CSDs that are superimposed on a common structural core. The analyses also revealed a previously unrecognised circumferential hydrophobic sash around the surface of the CD structure.

Conclusions: Dimerisation through the CSD of HP1-like proteins results in the simultaneous formation of a putative protein–protein interaction pit, providing a potential means of targeting CSD-containing proteins to particular chromatin sites.

Addresses: *Institute of Cell and Molecular Biology, The University of Edinburgh, Swann Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JR, Scotland. ‡MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, Scotland.

†N.P.C. and J.F.P. contributed equally to this work.

Correspondence: Paul J. McLaughlin
E-mail: P.McLaughlin@ed.ac.uk

Received: **24 February 2000**
Revised: **20 March 2000**
Accepted: **21 March 2000**

Published: **18 April 2000**

Current Biology 2000, 10:517–525

0960-9822/00/$ – see front matter

Background

Eukaryotic centromere regions have a distinct appearance and are generally transcriptionally inert. Transcription of genes placed within or close to these regions is frequently unstably silenced, resulting in mosaic phenotypes as typified by position-effect variegation. Similarly, regulated gene silencing contributes to the propagation of determined cellular states in development [1,2]. Pivotal components in the formation of these types of silent chromatin are the chromodomain-containing proteins.

The chromodomain is a motif shared by two conserved *Drosophila* proteins: HP1, a component of centromeric heterochromatin; and Polycomb, a protein required for homeotic gene repression [3]. Mammalian Polycomb is also required for normal development [2,4], and both mice and humans express three distinct forms of HP1, which show differential localisation, but some accumulation at pericentromeric heterochromatin [5–9]. The fission yeast *Schizosaccharomyces pombe* Swi6 protein is an HP1 homologue and associates with those chromosomal domains known to assemble transcriptionally silent chromatin [10–13]. Cells lacking *swi6* are viable but no longer form silent chromatin over the outer regions of centromeres [12]. These defective centromeres lead to a high frequency of chromosome segregation defects [10]. *Drosophila* HP1 mutations also affect chromosome segregation [14], and halving the dose of functional HP1 alleviates repression of genes embedded in pericentric heterochromatin while increased HP1 expression enhances their repression [15]. Similarly, elevated levels of mouse HP1β (also known as MoMOD1 or M31) in mouse cells cause dose-responsive silencing of centromeric transgenes [16]. These observations suggest a model in which the assembly of silent heterochromatin is driven by the concentration of several key components that associate and 'spread' from a nucleation site.

The Swi6 protein can spread over 3 kb of exogenous DNA to coat and silence a gene inserted within a fission yeast centromere [17]. The association of Swi6 with silent chromatin is dependent on the Rik1 and Clr4 proteins and their mutation also leads to silencing and chromosome segregation defects [12,18]. Clr4 is homologous to *Drosophila* and mammalian Suvar39 proteins and contains chromo and SET (another chromatin regulator motif found in Suvar39, Enhancer of Zeste and Trithorax) domains [19,20]. Mammalian Suvar39 coimmunoprecipitates with HP1β and is concentrated around centromeres [20].

Chromodomains are independent, globular domains and it is likely that they act as adaptor domains to organise multimeric complexes and facilitate assembly of extended silent chromatin domains. Two types of chromodomains have been described. Apart from an amino-terminal chromodomain (CD) [3,21], all HP1-like proteins, including Swi6, carry a carboxy-terminal chromo shadow domain (CSD) [21,22]. CSDs have only been found in proteins that also bear a CD, whereas CDs can exist in isolation. The roles of these two motifs appear to be distinct. The CD of Polycomb can provide chromatin-targeting activity [23,24]. The nuclear magnetic resonance (NMR) structure of the MoMOD1 CD suggested that it may act to mediate interactions with other proteins through an unusual hydrophobic groove near the amino terminus [25]. Although analyses in *Drosophila* suggest that the Polycomb CD mediates interactions with other Polycomb-group (PcG) proteins [26], no direct interactions with a CD have been reported. In contrast, a variety of factors are known to interact with HP1-like proteins through their CSD (reviewed in [27]). Comparisons of the CSD-interacting domain of these proteins identified a common PxVxL motif. Mutation of this motif in TIF1β and CAF1 p150 prevents association with CSDs [28–30]. Random peptide phage display analyses identified a consensus pentapeptide [PL][WRY]V[MIV][MLV] that is sufficient for interaction with the CSD of *Drosophila* HP1 [31]. Moreover, CSDs themselves self-interact and interact with other CSDs [28,32]. To gain further insight into how CSDs might mediate protein–protein interactions, the structure of the Swi6 CSD has been analysed after crystallisation. Here, the structure of the Swi6 CSD is presented and its implications discussed.

Results and discussion

Monomer fold

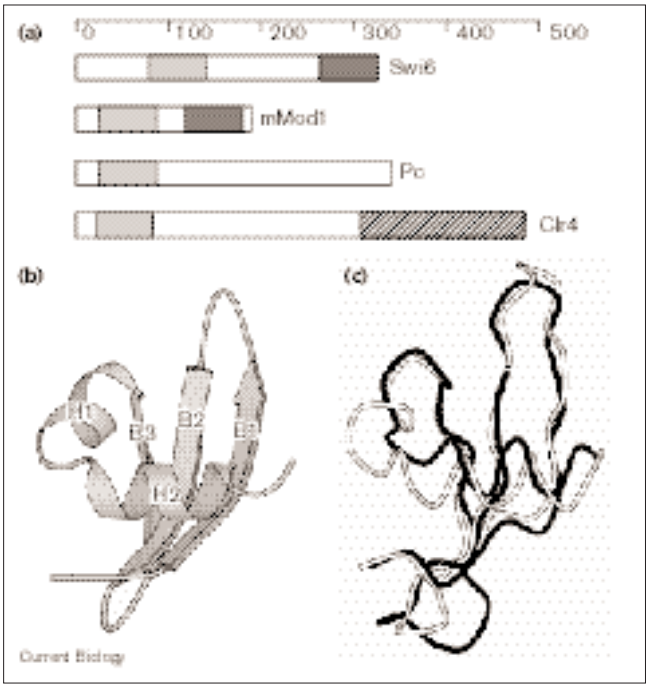
The structure of the CSD of the fission yeast Swi6 protein was solved by multiple isomorphous replacement and refined to 1.9 Å (see Supplementary material). The structure reveals that the CSD dimerises, with the dimer formed by contact between helices from different molecules related by a non-crystallographic twofold axis. Each monomer has dimensions of roughly 29 Å × 30 Å × 34 Å and is composed of three β strands that form an anti-parallel sheet, followed by two carboxy-terminal α helices (helix 1 and 2) that pack against the sheet (Figure 1b). This topology is similar to that of the CD, of which the MoMOD1 NMR structure is the only family member represented in the protein databank [25]. The two structures superimpose with a root mean square deviation (r.m.s.d.) of 1.8 Å for 45 Cα positions that correspond to most secondary structure elements. The exception is helix H1 (Figure 1c), which, interestingly, is found only in the CSD.

Using this structural correspondence, we constructed a sequence alignment of CD and CSD sequences (Figure 2a). Those residues conserved in both are shown in Figure 2b. They are all buried residues, pointing inwards from the β sheet and helix H2, and stack against each other to form a hydrophobic structural core that is common to both families. The most striking difference between the structures is the helix H1. The CSD family has an insertion of 2–3 residues in this region, corresponding to an α helix H1 that is not seen in the MoMOD1 CD structure. In addition, a proline (at position 311) that lies in the turn between helices H1 and H2 is absolutely conserved in the CSD family, but is only sometimes found in the CD family. Interestingly, this proline, and other conserved residues that are unique to the CSD family, are all found at the dimer interface.

Dimer structure

The dimer interface centres on helix H2, which interacts symmetrically with helix H2 of the other subunit (Figure 3a). The two helices cross at an angle of 35° and are closest at the amino-terminal end (Q312; Figure 3b). Thus, the helices splay further apart on progression towards the carboxyl terminus. Midway down helix H2,

Figure 1

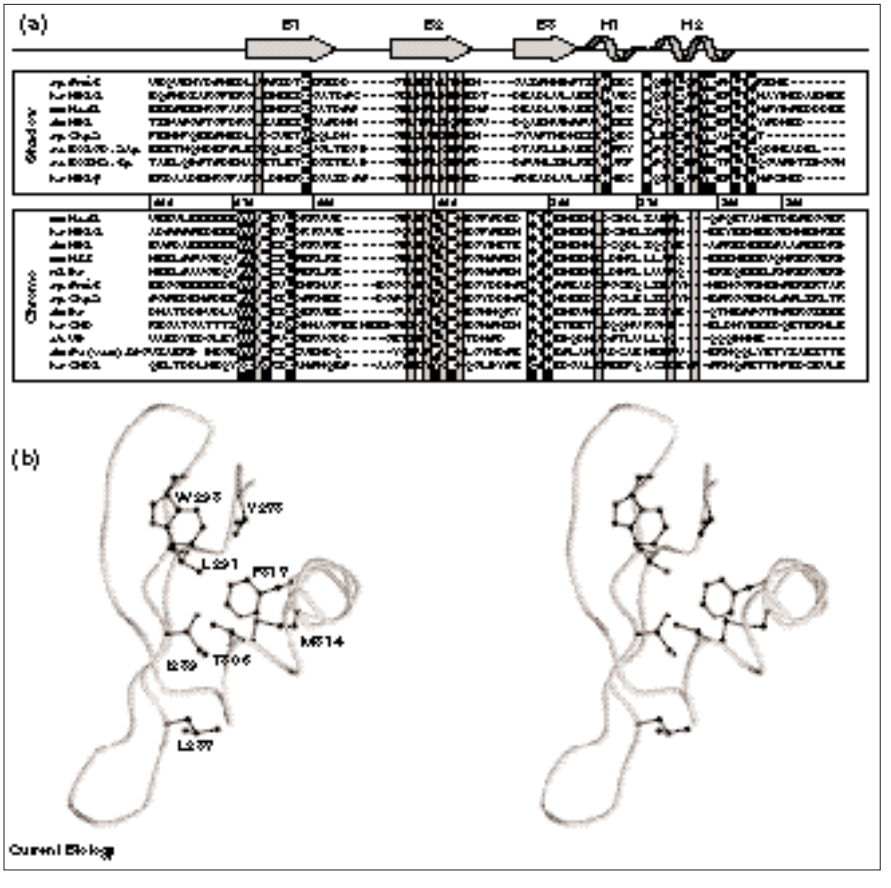


(a) Schematic diagram showing locations of CDs and CSDs in proteins involved in gene silencing and maintenance of heterochromatin. Light grey, CDs; dark grey, CSDs; hatched box, SET domain. Amino-acid numbers are indicated by a scale bar. Swi6 is *S. pombe* Swi6; mMod1 is *Mus musculus* MoMOD1 (identical to HP1β and M31); Pc is Polycomb from *D. melanogaster*; Clr4 is *S. pombe* Su(var) 3-9 homologue. **(b)** Monomer fold for Swi6 CSD residues 266–324. The secondary structure elements are numbered in order from amino to carboxyl terminus; B1, B2 and B3 are β strands, and H1 and H2 are α helices. The coordinates have been deposited in the protein databank, with pdb accession number 1e0b. **(c)** Superposition of the structures of Swi6 CSD (white) and MoMOD1 CD (black).

residues L315 and Y318 interact with the same residues in the other monomer. The dimerisation interface has more non-polar atoms than average for a protein–protein interface [33] (Figure 3c; 61% of the 1300 Å^2 of solvent-accessible surface that becomes buried in the dimer involves carbon atoms). Protein oligomers in which subunits are no longer soluble as independent molecules have, however, both significantly more non-polar and more extensive interfaces [34] than found in this dimer. Thus, the dimer is potentially in equilibrium with monomer, but dimerisation may be favoured. Consistent with this, in the monomer, both of the non-polar side chains (L315 and Y318) would be exposed to solvent; L315 in particular is exactly out of phase with conserved residues in the helix that form the buried hydrophobic core of the domain and is fully exposed to solvent, but would be buried in the dimer. In addition, intersubunit hydrogen bonds occur between residues at the top and bottom of helix H2 (Q312 and E319) and residues on helix H1 on the other subunit (N307 and S303). Figure 3b shows four direct hydrogen bonds and 10 further intersubunit contacts mediated by four water molecules. In particular, one water molecule forms a three-way contact between OH group of Y318 and Oε1 of E319 in helix H1 with Oγ of S303 in helix H2 of the other molecule (Figure 3d). Thus, a water molecule extends a network between residues that already form direct intersubunit contacts at the dimer interface.

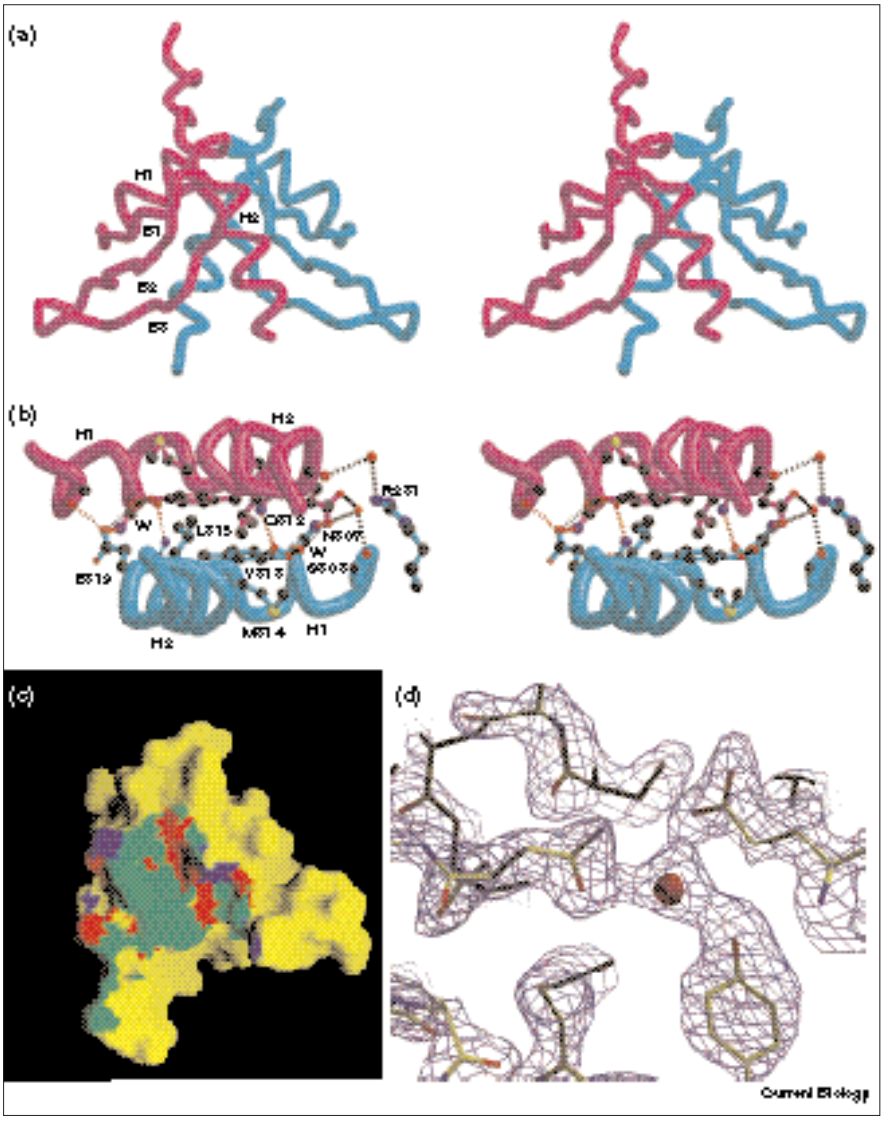
Both subunits of the dimer are very similar (r.m.s.d. main chain atoms, 0.36 Å, residues 266–320) even though they were built and refined independently. Nevertheless, there are some slight asymmetries between the structures seen for the monomers. At the amino terminus of the structure, residues 262–266 are visible in electron density maps in molecule A but not in B (Figure 3a). If molecule B had a similar configuration, residues 262–266 of the two molecules would clash within the dimer structure. This stretch, and the following five residues contain four negatively charged side chains that face the identical residues in the other subunit to form a negatively charged surface. None of the negatively charged side chains is absolutely conserved across the CSD family, however, suggesting that the charge and asymmetry is not a conserved feature of CSDs. On the other hand, molecule B has electron density from residues 322–324 while none is visible for molecule A. Curiously, at position 324, an aromatic (F) or branched non-polar side chain is conserved across the CSD family, without obvious explanation from the unliganded structure.

Figure 2



(a) Structure-based alignment of CSD and CD sequences. Secondary structure elements shown were derived from both mouse HP1β CD and *S. pombe* Swi6 CSD. Light grey, residues conserved in both families; dark grey, family-specific conservation. Standard protein naming has been used throughout. Species are identified by two letter abbreviations: sp, *S. pombe*; hs, *H. sapiens*; mm, *M. musculus*; ce, *Caenorhabditis elegans*; dm, *D. melanogaster*; xl, *Xenopus laevis*; at, *Arabidopsis thaliana*. Initial alignment within each family was calculated using ClustalW. Sequence numbering is relative to full-length Swi6 protein. **(b)** Stereographic view showing a backbone trace of Swi6 CSD appended with those residues that are conserved in both the CD and CSD families. The sequence numbering follows that in (a).

Figure 3



(a) Stereographic view of the dimer. Molecule A is shown coloured magenta, molecule B coloured cyan. The viewpoint is perpendicular to the dimer axis. **(b)** Stereographic view down the twofold axis showing inter-domain hydrogen bonding. For clarity, only helices H1 and H2 from each molecule are shown. Bonding residues are shown as sticks and labelled. A structural water molecule coordinated by the side chains of Y318, E319 and N307 is labelled W. **(c)** Nature of dimer contact surface. A monomer in the same orientation as the cyan monomer in (a) is shown in yellow. Those atoms that contact the other subunit in the dimer and are carbons are shown in green while nitrogens are shown in blue and oxygen in red. **(d)** A simulated annealing map at 1.9 A is shown at 1.5 σ, showing density around the water molecule in (b) that forms three hydrogen bonds between subunits. The map coefficients are 2Fo–Fc.

In vitro association of CSD proteins

The crystal structure clearly indicates that the Swi6 CSD dimerises, and that residues such as L315, which are closely aligned in the dimerisation interface, may be important for the dimerisation. The ability of the CSD to self-dimerise *in vitro* was tested. Purified glutathione-S-transferase (GST) fusion proteins of the CD, CSD, and CSD with L315 mutated to Asp (CSD L315D), were tested for dimerisation with the non-tagged CSD in crude bacterial lysate. GST fusion proteins and their associated polypeptides were purified by incubation of these mixtures with glutathione–agarose beads. Initially, no binding of the CSD to any of the GST fusion proteins could be detected, even using a thirtyfold molar excess of the CSD in the binding reactions. This suggested that the dimerisation of GST–CSD molecules was very stable, and refractory to exchange with CSD molecules, or that, under the binding conditions used, dimers could not form.

To test these possibilities, urea denaturation was performed on mixtures of GST fusion proteins and the bacterial extract containing Swi6 CSD to disrupt pre-formed dimers. This was followed by renaturation of the polypeptides. After binding of GST fusion proteins to glutathione beads, and extensive washing, bound polypeptides were eluted and visualised by Coomassie staining of polypeptides resolved by SDS–PAGE. From Figure 4, it is clear that the CSD was bound only to the GST–CSD fusion protein. There was no apparent binding under these conditions to the GST–CD protein, or to CSD L315D. The western blot analysis in Figure 4 confirms that the CSD protein is bound specifically to GST–CSD, and not to the CD or mutated

CSD, and suggests that there is a 1:1 binding of CSD to GST–CSD when the CSD is present in excess.

These experiments therefore suggest that dimerisation of CSD proteins can occur *in vitro*, and that the dimer is stable. In addition, mutation of a single amino acid at the dimerisation interface (L315) is sufficient to prevent dimerisation in this assay.

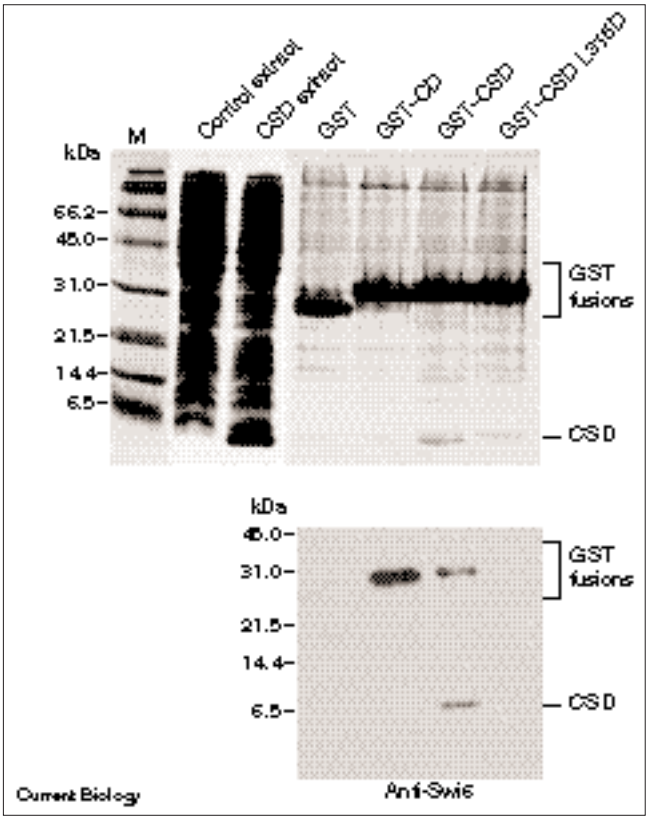
Potential binding sites for other proteins

In the MoMOD1 CD structure, a hydrophobic groove has been noted, and proposed to form a site for protein interaction [25]. Many proteins that interact with CSD proteins bear a conserved hydrophobic pentapeptide motif [28–31]. It has been suggested that the hydrophobic groove is also conserved in the CSD, and that the peptide binds within this groove [31]. When we compared the structure of the Swi6 CSD with that of the MoMOD1 CD, we found, however, that amino acids just amino-terminal to the first β sheet of the CSD occupy and occlude this groove in the CSD (Figure 5a). It is therefore unlikely that this groove is a conserved protein-interaction surface within CSDs. We note that the equivalent sequence within the MoMOD1 CD consists of a polyglutamic acid stretch, which is not likely to block the groove, perhaps leaving it accessible for potential interactions.

Nevertheless, we noted that, as the H2 helices of the CSD dimer diverge at the dimer interface towards the carboxyl terminus of the molecules, a pit with a non-polar base develops. Roughly halfway down the helices, residues corresponding to L315 and Y318 in all CSDs are branched or aromatic, and can pack the space between the diverging helices. These amino acids, contributed from both molecules, form the base of the non-polar pit.

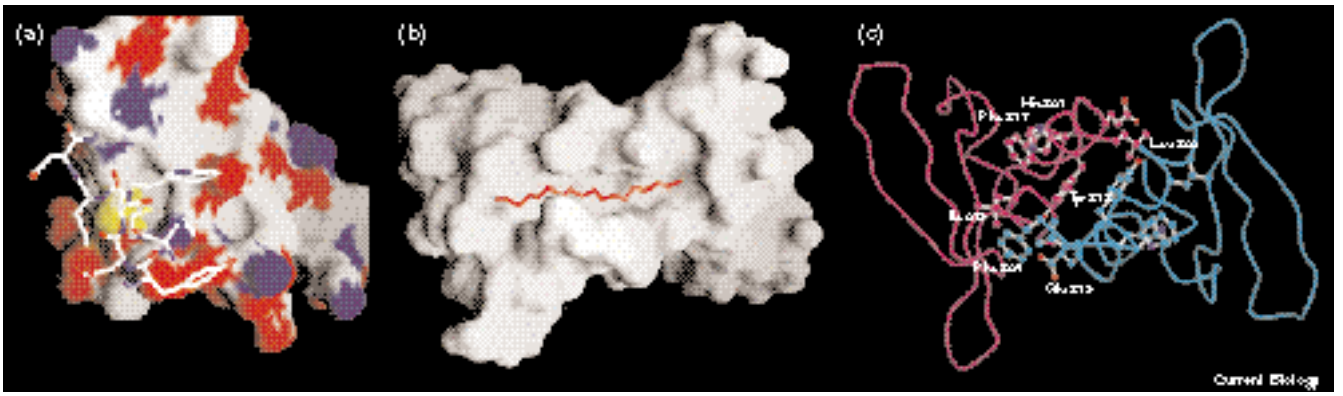
Many of the proteins that interact with CSDs are characterised by the conservation of a pentapeptide sequence of consensus PxVxL, and recent peptide phage display experiments have shown that peptides of [PL][WRY]V[MIL][MLV] sequence can bind the CSD of *Drosophila* HP1 [31]. To test whether pentapeptide sequences of this type could fill this non-polar pit on the CSD dimer surface, five unique examples of this sequence were retrieved from the whole-protein structural databank. Four of these were solvent accessible, and all of the peptides were in an extended conformation. Thus, there are precedents for such non-polar sequences being presented at the surface of a globular protein. Next, as all examples were extended structures, we used these as rulers to try to dismiss the pit as a potential binding site for the pentapeptide motifs. In fact, these peptides (with their full side chains) can easily be accommodated into the pit of the CSD dimer to closely pack its interior. In addition, there are some other attractive features for this site. As the peptide binds across a twofold axis, the model predicts one pentapeptide binding site per CSD dimer. As peptides are asymmetric molecules that do not have twofold axes, a corollary is that the peptide should induce asymmetry in the site. We have, however, already noted asymmetry at the carboxyl terminus of the dimer structure in terms of electron density for two more residues, one of which is conserved in type across CSDs (F324). Therefore, asymmetry may already be built into the unliganded dimer structure to provide an asymmetric lock for an asymmetric key. In addition, the binding site may explain the puzzling conservation of F324. It is close enough to the pit to move to complete binding of the hydrophobic target sequence and may act as a 'cover' over the bound peptide. Another residue that is conserved exclusively within the CSD family, and lies at each end of the pit is I279 (Figure 5c). The conservation of both F324 and I279 cannot be explained by the fold or dimerisation of the unliganded structure. It is

Figure 4



In vitro dimerisation of the Swi6 CSD and its dependence on L315. Crude bacterial lysates, with or without expression of non-tagged CSD, were mixed with purified GST fusion proteins. The GST fusion proteins and associated polypeptides were separated by PAGE and visualised by Coomassie staining after urea denaturation, refolding and purification on glutathione–agarose. CSD bound only to GST–CSD, and binding was abrogated by mutation of L315 to Asp (GST–CSD L315D). The bottom panel shows western blot analysis of the samples using anti-Swi6 antibody. M, protein markers.

Figure 5



(a) Molecular surface of residues 22–69 of the CD of MoMOD1. White, carbon; red, oxygen; blue, nitrogen; yellow, sulphur. The residues of Swi6 CSD have been superimposed by a least-squares algorithm. The view is centred on a deep hydrophobic groove noted by Ball *et al.* [25] in the CD structure, and proposed to be conserved in CSDs [31]. In the CSD of Swi6, amino-terminal residues can be seen to cover this groove. **(b)** A surface representation is shown of a dimer in the same orientation as in (c), looking down the dimer axis from the carboxyl terminus of the helices. A non-polar pit is delineated by a pentapeptide carbon backbone (red). The figure was generated by GRASP. **(c)** A backbone trace of the dimer in a similar orientation as that in (b). The side chains of residues that line the non-polar pit are shown. The trace was generated using MOLSCRIPT, followed by RASTER3D.

interesting that both conserved residues cluster to a site suitable for binding an extended, non-polar pentapeptide.

Chromodomain family fold

All conserved surface residues in CSD-containing proteins map to the dimer interface. These are concentrated around helices H1 and H2 (Figure 2a). To understand why CDs do not dimerise [25], we mapped the MoMOD1 CD monomer onto the Swi6 CSD dimer. There are few clashes and, at first sight, the artificial model seems to pack quite well. The crucial non-polar residues Y318 and L315 that are at the dimerisation interface are not conserved and are often uncompensated, charged residues in the CD family, however. By contrast, in CSDs, Y318 is absolutely conserved and L315 is either L or I. Smaller side chains could not pack the space between the helices, and the formation of a stable base to the potential peptide pit could further constrain the types of side chains at these positions. Also, in the CD family, several hydrogen-bonding residues that are absolutely conserved in the CSD family are not present; these are E319 in H2 and N307 in H1 (the latter corresponding to the region of insertion/deletion in H1). All of these differences between the CD and CSD suggest that dimer formation is preferred in CSDs, whereas CDs are monomeric.

Conserved chromodomain surface

Comparison of the sequences and structures between the two families revealed that the CD family displays more conservation in the three β strands than the CSD. When these regions of conservation were mapped onto the surface of the MoMOD1 CD, it was clear that they lined up across the face of the sheet to form a non-polar stripe (Figure 6b) with contributions from all three strands, which wind around the structure, reminiscent of a sash. This striking feature might therefore represent a conserved surface feature that is recognised by a binding partner. Interestingly, a conservative substitution in a residue that maps to the sash in HP1 from *D. melanogaster* results in loss of silencing [24]. This mutation corresponds to a Y→F change at position 270 in Figure 2a. The CD structure predicts that this mutant protein is still folded and other analysis showed that it is targeted to chromatin normally [24]. Such a mutation may perturb binding of an HP1 CD partner. To date, there are no known direct protein interactions that occur solely through CDs. As the hydrophobic sash runs circumferentially around the surface of the CD, it is likely that more than one protein is required for interaction, and this may explain why to date no CD-interacting proteins have been identified by two-hybrid approaches. In addition, peptide phage display analyses with the CD of HP1 failed to provide a consensus peptide-interaction motif [31].

Conclusions

The CSD structure has revealed that, although similar, the CD and CSD have distinct structures and binding motifs based on a common structural core. Homologies between the two structures reside in their hydrophobic core, generated by three β strands and helix H2. Notably, the structural analysis suggests that the CSD provides at least two modes of protein interaction: first, the formation of stable dimers and, second, dimerisation generates an interaction pit that may allow docking of molecules with a hydrophobic pentapeptide sequence. It is possible that interactions with pentapeptide proteins provide functions

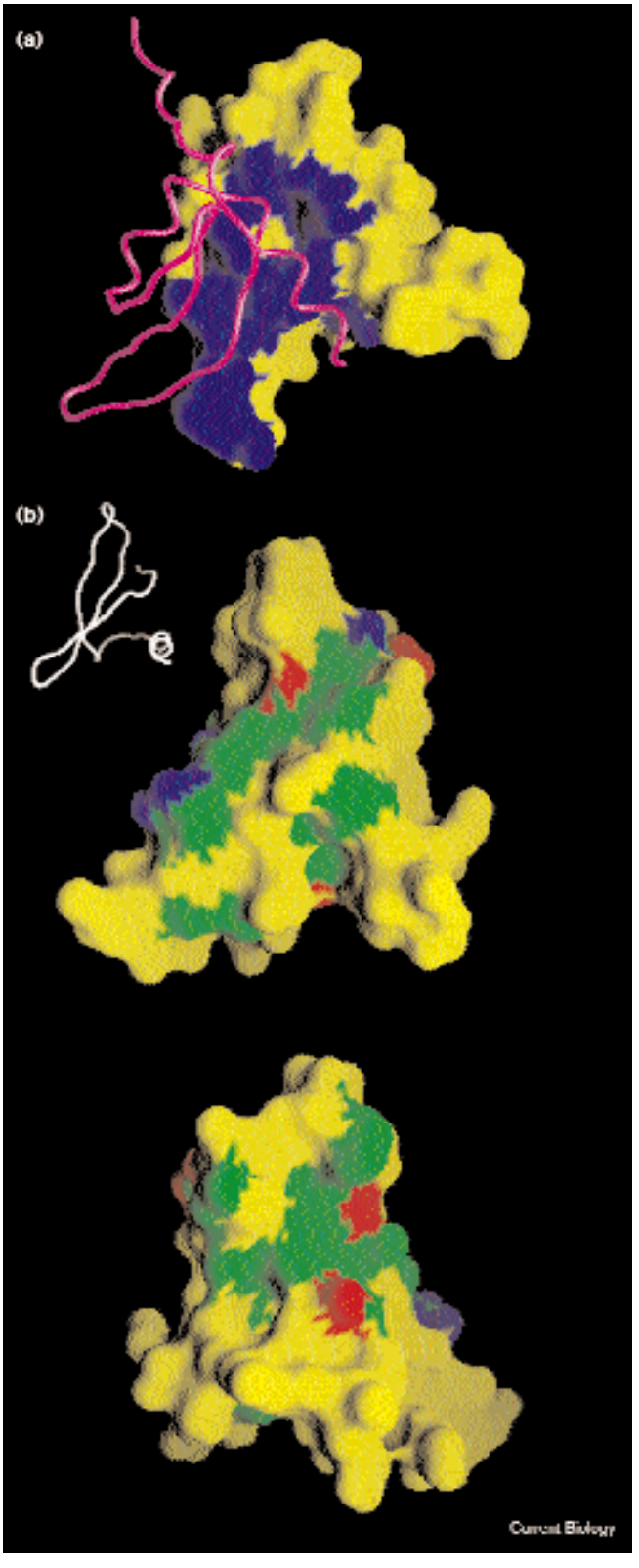


Figure 6

(a) Conserved dimer interface. Residues that are conserved uniquely in CSD are mapped (blue) onto the surface of Swi6 CSD structure. The view is similar to the cyan molecule in Figure 3a. The second subunit is shown in magenta for reference. **(b)** Conserved surface residues in the CD (MoMOD1). Those residues uniquely conserved in the CD family are mapped onto the surface according to atom type (green, carbon; blue, nitrogen; red, oxygen). Note that the conserved residues form a continuous sash around the molecule. The lower view is rotated 180° around the Y axis. A backbone trace is shown for the top view.

such as targeting CSD-containing proteins to particular sites of chromatin. The possibility of heterodimerisation of CSD proteins may further contribute to the diversity of proteins that interact with the interaction pit.

Materials and methods

Plasmid constructs

Swi6 CSD (amino acids 261–328) was amplified from pAL2 [35] using oligonucleotides incorporating *Bam*HI and *Eco*RI restriction sites and a stop codon, and subcloned into a pET-based expression vector (pMW172). The expressed CSD protein therefore had three amino acids contributed from the vector sequence and reads, from the amino terminus, MGSKQ… The numbering system used for Swi6 CSD starts from K261 to E328.

GST fusion constructs were generated by PCR of the Swi6 CD (amino acids 80–133, GST–CD), the CSD (as above), or CSD with L315 mutated to Asp (GST–CSD L315D), with wild-type or mutagenic primers bearing *Bam*HI and *Eco*RI sites. Isolated fragments were cloned in frame with GST into pGEX-KG [36]. Integrity of cloned DNA was verified by sequence analysis.

Expression and purification of recombinant proteins

Crude CSD bacterial extract for the *in vitro* binding assays was prepared by extraction of BL21s bearing CSD-expressing pMW172. After a 3 h induction with 0.1 mM IPTG at 34°C, cells were harvested, washed, and lysed by sonication in 300 mM NaCl, 25 mM Tris pH 8, 1 mM EDTA, 1 mM EGTA, 10% (v/v) glycerol and protease inhibitors. Extracts were clarified by centrifugation at 13,000 × *g* at 4°C for 15 min, and stored at −70°C.

GST and GST fusion proteins were expressed in BL21 (DE3) bacteria. Overnight cultures were diluted 1 in 100, and grown at 34°C to an OD_{600} of 0.6. Protein expression was induced by addition of 0.1 mM IPTG for 3 h. Cells were collected, washed with PBS and extracted as described [37]. After binding to glutathione–agarose, and washing, bound proteins were eluted with 5 mM glutathione, dialysed against PBS containing 10% (v/v) glycerol and stored at −70°C.

Purification and crystallisation

For crystallisation, soluble CSD was purified by fractionation using two ion-exchange steps from extracts of IPTG-induced BL21 (DE3) *Escherichia coli* carrying pMW172–CSD.

Proteins in 50 mM NaCl, 25 mM Tris pH 7.0, 1 mM EDTA were eluted from a DE52 Sepharose anion exchange column across a 50–500 mM NaCl gradient. Fractions containing Swi6 CSD were pooled and dialysed against buffer A (25 mM NaCl, 25 mM Tris pH 6.0, 1 mM EDTA). Following centrifugation, the sample was concentrated to 5 ml using a Vivaspin low MW cut-off spin concentrator. After loading onto a CM52 Sepharose column, CSD was eluted in buffer A. Purified CSD was concentrated to 15 mg/ml and crystallised by vapour diffusion. Almond-shaped crystals of approximately 0.5 mm × 0.3 mm × 0.2 mm were grown in 20% (w/v) PEG 4000, 0.2 M sodium acetate, 0.1 M Tris HCl pH 8.5 at 4°C.

Data collection

Crystals were cooled in liquid nitrogen in their mother liquor plus 4% (v/v) glycerol. Diffraction data to 1.9 Å was collected using a CCD (Mar) scanner at the SRS Daresbury beamline PX9.6 with a wavelength

of 0.87 Å. The native data was indexed using the program MOSFLM. Heavy atom derivatives were obtained by soaking native crystals in 1 mM EMTS for 26 h, 1 mM K_2PtCl_4 for 24 h and a double soaked crystal in 1 mM EMTS and 0.5 mM K $Au(CN)_2$ for 26 h. All soaks were carried out at 4°C. Diffraction data from the derivative crystals was collected on a home source generator.

Sequence alignments

ClustalW was used to align CDs and CSDs families separately. The validity of each such alignment was checked against the representative structure (MoMod1 CD and Swi6 CSD). Then, each family was aligned to the other by noting the correspondence between conserved positions common to both families and equivalent positions in regular secondary structure in both structures.

In vitro *binding assays*

Recombinant GST fusion proteins (40 μg) were mixed with 100 μl crude bacterial extract (containing approximately 2 mg/ml CSD). Proteins were denatured and renatured by dialysis against 8 M urea, followed by 6 M, 4 M, 2 M, 1 M and 0.5 M urea, and finally glutathione–agarose binding buffer with no urea. Glutathione–agarose beads (50 μl) were added, incubated for 30 min at 4°C, then the beads washed three times with 1 ml binding buffer. Bound proteins were eluted with sample buffer and heating, before loading on a 16% Prosieve 50 (FMC) SDS–polyacrylamide gel and electrophoresis in Tricine buffer. Resolved polypeptides were visualised by Coomassie staining. Western blot analysis was performed by loading 1/200 volume of the sample used for the Coomassie-stained gel, and probing the blot with affinity-purified anti-Swi6 antibodies [10].

Supplementary material

Supplementary material including additional methodological detail is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We thank M. Harding for help in constructing the polyethylene glycol molecule and Catriona Mowbray for her initial work on CD protein expression. This work was funded by core support from the Medical Research Council of Great Britain to R.C.A.; by an MRC postgraduate studentship to N.P.C.; and Wellcome funding to P.J.M. We are grateful for beam time at SRS, Daresbury.

References

1. Paro R, Harte PJ: **The role of Polycomb group and trithorax group chromatin complexes in the maintenance of determined cell states.** In *Epigenetic Mechanisms of Gene Regulation*. Edited by Russo E. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1996:507-528.
2. Core N, Bel S, Gaunt SJ, Aurrand-Lions M, Pearce J, Fisher A, Djabali M: **Altered cellular proliferation and mesoderm patterning in Polycomb-M33 deficient mice.** *Development* 1997, **124:**721-729.
3. Paro R, Hogness DS: **The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*.** *Proc Natl Acad Sci USA* 1991, **88:**263-267.
4. Katoh-Fukui Y, Tsuchiya R, Shiroishi T, Nakahara Y, Hashimoto N, Noguchi K, Higashinakagawa T: **Male-to-female sex reversal in M33 mutant mice.** *Nature* 1998, **393:**688-692.
5. Nicol L, Jeppesen P: **Human autoimmune sera recognize a conserved 26 kD protein associated with mammalian heterochromatin that is homologous to heterochromatin protein 1 of *Drosophila*.** *Chromosome Res* 1994, **2:**245-253.
6. Wreggett KA, Hill F, James PS, Hutchings A, Butcher GW, Singh PB: **A mammalian homolog of *Drosophila* heterochromatin protein-1 (HP1) is a component of constitutive heterochromatin.** *Cytogenet Cell Genet* 1994, **66:**99-103.
7. Horsley D, Hutchings A, Butcher GW, Singh PB: **M32, a murine homologue of *Drosophila* heterochromatin protein 1 (HP1), localises to euchromatin within interphase nuclei and is largely excluded from constitutive heterochromatin.** *Cytogenet Cell Genet* 1996, **73:**308-311.
8. Furuta K, Chan EKL, Kiyosawa K, Reimer G, Luderschmidt C, Tan EM: **Heterochromatin protein HP1(Hs beta) (p25 beta) and its localization with centromeres in mitosis.** *Chromosoma* 1997, **106:**11-19.
9. Minc E, Allory V, Worman H J, Courvalin JC, Buendia B: **Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells.** *Chromosoma* 1999, **108:**220-234.
10. Ekwall K, Javerzat JP, Lorentz A, Schmidt H, Cranston G, Allshire RC: **The chromodomain protein swi6 — a key component at fission yeast centromeres.** *Science* 1995, **269:**1429-1431.
11. Allshire RC, Javerzat JP, Redhead NJ, Cranston G: **Position effect variegation at fission yeast centromeres.** *Cell* 1994, **76:**157-169.
12. Allshire RC, Nimmo ER, Ekwall K, Javerzat JP, Cranston G: **Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation.** *Genes Dev* 1995, **9:**218-233.
13. Nimmo ER, Cranston G, Allshire RC: **Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes.** *EMBO J* 1994, **13**:3801-3811.
14. Kellum R, Alberts BM: **Heterochromatin protein-1 is required for correct chromosome segregation in *Drosophila* embryos.** *J Cell Sci* 1995, **108**:1419-1431.
15. Eissenberg JC, Morris GD, Reuter G, Hartnett T: **The heterochromatin-associated protein HP-1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation.** *Genetics* 1992, **131:**345-352.
16. Festenstein R, Sharghi-Namini S, Fox M, Roderick K, Tolaini M, Norton T, *et al.*: **Heterochromatin protein 1 modifies mammalian PEV in a dose- and chromosomal-context-dependent manner.** *Nat Genet* 1999, **23:**457-461.
17. Partridge JF, Borgstrom B, Allshire RC: **Distinct protein interaction domains and protein spreading in a complex centromere.** *Genes Dev* 2000, **14:**783-791.
18. Ekwall K, Nimmo ER, Javerzat JP, Borgstrom B, Egel R, Cranston G, Allshire RC: **Mutations in the fission yeast silencing factors clr4(+) and rik1(+) disrupt the localisation of the chromo domain protein swi6p and impair centromere function.** *J Cell Sci* 1996, **109:**2637-2648.
19. Ivanova AV, Bonaduce MJ, Ivanov SV, Klar AJS: **The chromo and set domains of the clr4 protein are essential for silencing in fission yeast.** *Nat Genet* 1998, **19:**192-195.
20. Aagaard L, Laible G, Selenko P, Schmid M, Dorn R, Schotta G, *et al.*: **Functional mammalian homologues of the *Drosophila* PEV-modifier su(var)3–9 encode centromere-associated proteins which complex with the heterochromatin component m31.** *EMBO J* 1998, **18:**1923-1938.
21. Koonin EV, Zhou S, Lucchesi JC: **The chromo superfamily: new members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin.** *Nucl Acids Res* 1995, **23:**4229-4233.
22. Aasland R, Stewart AF: **The chromo shadow domain, a 2nd chromo domain in heterochromatin-binding protein-1, HP1.** *Nucl Acids Res* 1995, **23:**3168-3173.
23. Messmer S, Franke A, Paro R: **Analysis of the functional role of the Polycomb chromodomain in *Drosophila melanogaster*.** *Genes Dev* 1992, **6:**1241-1254.
24. Platero JS, Hartnett T, Eissenberg JC: **Functional-analysis of the chromo domain of HP1.** *EMBO J* 1995, **14:**3977-3986.
25. Ball LJ, Murzina NV, Broadhurst RW, Raine ARC, Archer SJ, Stott FJ, *et al.*: **Structure of the chromatin binding (chromo) domain from mouse modifier protein1.** *EMBO J* 1997, **16:**2473-2481.
26. Strutt H, Paro R: **The Polycomb group protein complex of *Drosophila melanogaster* has different compositions at different target genes.** *Mol Cell Biol* 1997, **17:**6773-6783.
27. Jones DO, Cowell IG, Singh PB: **Mammalian chromodomain proteins: their role in genome organisation and expression.** *Bioessays* 2000, **22:**124-137.
28. LeDouarin B, Nielsen AL, Garnier JM, Ichinose H, Jeanmougin F, Losson R, Chambon P: **A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors.** *EMBO J* 1996, **15:**6701-6715.
29. Murzina N, Verreault A, Laue E, Stillman B: **Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins.** *Mol Cell* 1999, **4:**529-540.
30. Nielsen AL, Ortiz JA, You J, OuladAbdelghani M, Khechumian R, Gansmuller A, *et al.*: **Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family.** *EMBO J* 1999, **18:**6385-6395.
31. Smothers JF, Henikoff S: **The HP1 chromo shadow domain binds a consensus peptide pentamer**. *Curr Biol* 2000, **10:**27-30.

32. Ye Q, Callebaut I, Pezham A, Courvalin J, Worman HJ: **Domain-specific interactions of human HP1-type chromodomain proteins and inner nuclear membrane protein LBR.** *J Biol Chem* 1997, **272:**14983-14989.
33. LoConte L, Chothia C, Janin J: **The atomic structure of protein-protein recognition sites.** *J Mol Biol* 1999, **285:**2177-2198.
34. Janin J, Miller S, Chothia C: Surface, **Subunit interfaces and interior of oligomeric proteins.** *J Mol Biol* 1988, **204:**155-164.
35. Lorentz A, Ostermann K, Fleck O, Schmidt H: **Switching gene Swi6, involved in repression of silent mating-type loci in fission yeast, encodes a homolog of chromatin-associated proteins from *Drosophila* and mammals.** *Gene* 1994, **143:**139-143.
36. Guan KL, Dixon JE: **Eukaryotic proteins expressed in *Escherichia coli* — an improved thrombin cleavage and purification procedure of fusion proteins with glutathione-s-transferase.** *Anal Bioch* 1991, **192:**262-267.
37. Smith DB, Johnson KS: **Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase.** *Gene* 1988, **67:**31-40.

Because *Current Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cub – for further information, see the explanation on the contents page.