Analysis of the Structure and Function of the Transcriptional Coregulator HOP[†,‡]

Hyun Kook,[§,#] Wendy W. Yung,[||,#] Raina J. Simpson,[||,#] Hae Jin Kee,[§] Sera Shin,[§] Jason A. Lowry,[||] Fionna E. Loughlin,[||] Zhan Yin,[⊥] Jonathan A. Epstein,[⊥] and Joel P. Mackay*[,||]

Medical Research Center for Gene Regulation, Chonnam National University Medical School, Gwangju, 501-746, South Korea, School of Molecular and Microbial Biosciences, University of Sydney, NSW 2006, Australia, and Cardiovascular Division, University of Pennsylvania Health System, Philadelphia, Pennsylvania 19104

Received April 1, 2006; Revised Manuscript Received July 3, 2006

ABSTRACT: Homeodomain-only protein (HOP) is an 8-kDa transcriptional corepressor that is essential for the normal development of the mammalian heart. Previous studies have shown that HOP, which consists entirely of a putative homeodomain, acts downstream of Nkx2.5 and associates with the serum response factor (SRF), repressing transcription from SRF-responsive genes. HOP is also able to recruit histone deacetylase (HDAC) activity, consistent with its ability to repress transcription. Unlike other classic homeodomain proteins, HOP does not appear to interact with DNA, although it has been unclear if this is because of an overall divergent structure or because of specific amino acid differences between HOP and other homeodomains. To work toward an understanding of HOP function, we have determined the 3D structure of full-length HOP and used a range of biochemical assays to define the parts of the protein that are functionally important for its repression activity. We show that HOP forms a classical homeodomain fold but that it cannot recognize double stranded DNA, a result that emphasizes the importance of caution in predicting protein function from sequence homology alone. We also demonstrate that two distinct regions on the surface of HOP are required for its ability to repress an SRF-driven reporter gene, and it is likely that these motifs direct interactions between HOP and partner proteins such as SRF- and HDAC-containing complexes. Our results demonstrate that the homeodomain fold has been co-opted during evolution for functions other than sequence-specific DNA binding and suggest that HOP functions as an adaptor protein to mediate transcriptional repression.

HOP[1] is a recently described protein that contains significant sequence homology to homeodomain-containing transcription factors. The homeodomain is a 60 amino acid motif that has been conserved across millions of years of evolution and is found in a large number of DNA-binding proteins, including Hox proteins that specify position along the anterior−posterior axis of the embryo in all metazoa. All previously described homeodomain-containing proteins include the 60 amino acid DNA-binding motif in addition to other domains that mediate transcriptional activation, repression, or protein−protein interactions. In contrast, HOP is composed almost entirely of a homeodomain-like sequence.

Primary sequence analysis indicates that HOP is closely related to members of the homeodomain family. The amino acid sequence of HOP is 40−45% identical to the homeodomains of Pax6 and Gsc (Figure 1). However, HOP has a number of unusual characteristics at the sequence level that distinguish it from other known homeodomains. In particular, it lacks several key residues that are conserved among all known homeodomains and are known to mediate critical contacts with DNA. In the absence of these conserved residues, HOP has been predicted to lack DNA-binding capacity. However, the observation that HOP does not interact with known homeodomain DNA-binding sequences could also be due to divergent structure because it has not been previously resolved whether HOP folds like other homeodomain proteins and adopts a similar 3D structure. Rather than binding to DNA, HOP is thought to function by interacting with other nuclear proteins, including serum response factor (SRF), to negatively modulate transcription during cardiac development. We and others recently demonstrated that disruption of the HOP gene in either mice or zebrafish results in severe cardiac abnormalities (*1*, *2*), and a series of in vitro and in vivo studies established that HOP acts downstream of the central cardiac transcription factor Nkx2−5 (*1*−*3*). It was also demonstrated that HOP can form a complex that includes SRF, inhibiting the ability of the latter to bind to DNA, and can repress the expression of reporter genes driven by SRF-dependent cardiac-specific promoters.

We have shown more recently that the transcriptional repression induced by HOP is blocked by the addition of

[†] This work was supported by the Australian Research Council, the National Institutes of Health, and a Korea Research Foundation Grant funded by the Korean Government (MOEHRD, KRF-2004-202-E00041). J.P.M. is an NHMRC Senior Research Fellow. H.J.K. was supported by Korea Research Foundation Grant KRF-2004-C00141.

[‡] The structure has been deposited in the Protein Data Bank under acecession code 2HI3. The BioMagResBank code for the NMR data is 7203.

* To whom correspondence should be addressed. Phone: +61-2-9351-3906. Fax: +61-2-9351-4726. E-mail: j.mackay@mmb.usyd.edu.au.

[§] Chonnam National University Medical School.

[||] University of Sydney.

[⊥] University of Pennsylvania Health System.

[#] These authors contributed equally to this work.

[1] Abbreviations: HOP, homeodomain only protein; SRF, serum response factor.

10.1021/bi060641s CCC: $33.50
Published on Web 08/11/2006

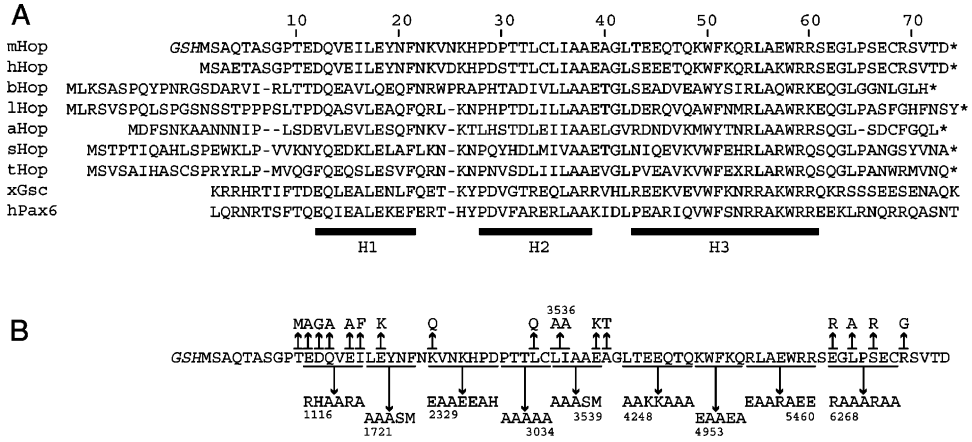

FIGURE 1: Amino acid sequence of HOP, HOP orthologues, and HOP mutants discussed in this study. (A) Sequences of HOP proteins. The *N*-terminal GSH shown for mHOP is a cloning artifact. The sequences of HOP-related proteins from several invertebrates are shown; bHOP, honey bee (*Apis mellifera*; EST acc# BI515127); lHOP, migratory locust (*Locusta migratoria*; EST acc# CO862286); aHOP, pea aphid (*Acyrthosiphon pisum*; EST acc# CN755981); sHOP, Mediterranean checkered scorpion (*Mesobuthus gibbosus*; EST acc# CB333861); and tHOP, black-legged tick (*Ixodes scapularis*; EST acc# DN971550). A sequence comparison with the two most closely related homeodomains, those from *Xenopus goosecoid* and human Pax6, is also shown. The numbering of full-length HOP is indicated above the sequence. (B) Representation of HOP mutants constructed for this work. Multiple-residue mutants are referred to by the two residues at either end of the mutated range (e.g., 1116 for a mutation of residues 11−16).

histone deacetylase (HDAC) inhibitors and that a HOP-containing complex immunoprecipitated from COS cells displays considerable HDAC activity (*4*). These results point to an additional mechanism through which HOP is able to repress the activity of certain SRF-dependent genes.

In this article, we describe the 3D solution structure of the divergent HOP homeodomain. We also provide evidence that unlike all other homeodomains characterized to date, HOP is unable to recognize double-stranded DNA, and we go on to map the sequence motifs within HOP that are important for its ability to mediate the repression of SRF-driven genes.

EXPERIMENTAL PROCEDURES

Preparation of HOP. Full-length HOP was cloned by PCR amplification from a cDNA template, ligated into the pET15b expression vector (Novagen) and expressed as a His-tagged fusion protein in *E. coli* BL21 cells using Luria broth. Cell pellets were lysed in a buffer containing 20 mM Tris (pH 8.0), 150 mM NaCl, 1.4 mM β-mercaptoethanol, and 0.5 mM phenylmethylsulfonylfluoride. The fusion tag was cleaved using thrombin (3 h at 37 °C) and further purified using reversed phase HPLC (C18). The final product comprised full-length HOP, carrying the additional three *N*-terminal amino acids (GSH; numbered −3 to −1, respectively) derived from the thrombin cleavage site. The identity of the protein was confirmed using electrospray mass spectrometry. ^{15}N- and ^{15}N,^{13}C-labeled HOP were prepared using the procedure of Cai et al. (*5*) and purified as described above.

Circular Dichroism Spectropolarimetry. Samples of HOP or HOP mutants were prepared by dissolving lyophilized HOP in a buffer containing 50 mM sodium phosphate and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP). The pH was adjusted to 6.8 using 0.1 M NaOH. CD spectra were recorded at 25 °C on a Jasco J-720 spectropolarimeter, equipped with a Neslab RTE−111 temperature controller. CD data were collected over the wavelength range 184−260 nm with a resolution of 0.5 nm, a bandwidth of 1 nm, and a response time of 1 s using a 1-mm path length cell. Final spectra were the sum of three scans accumulated at a speed of 20 nm/min and were baseline corrected and smoothed using a three-point running average algorithm.

Sedimentation Equilibrium. Sedimentation equilibrium experiments were carried out at 25 °C using a Beckman Optima XL-A ultracentrifuge equipped with an AnTi-60 rotor. Absorbance (280 and 360 nm) versus radius scans (0.001-cm increments) were collected at 3-h intervals until the samples had reached equilibrium. Data were recorded at three different speeds (30 000, 35 000, and 42 000 rpm) and two loading concentrations (10 and 20 μM). Analysis of the data was carried out using NONLIN software (*6*). The density of the solvent and the partial specific volume of the peptide (adjusted to include the contribution of Zn(II); $p = 0.719$ g/mL) were estimated using Sednterp (*7*). Final parameters were determined by a nonlinear least-squares fit to models incorporating homooligomerization, and an examination of residuals was used to determine the goodness of fit.

NMR Spectroscopy. For NMR experiments, purified HOP was dissolved in a buffer containing 50 mM sodium phosphate and 0.5 mM TCEP (pH 6.5−6.9) to final concentrations of up to ∼1.5 mM. NMR samples also contained 20 μM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) as a chemical shift reference and 5% (v/v) D_2O. Spectra were recorded at 25 °C on a Bruker DRX600 spectrometer. Samples of HOP mutants were prepared similarly. All homonuclear 2D data were collected and analyzed as described (*8*). Mixing times were 70 and 150 ms for TOCSY and NOESY spectra, respectively. ^{15}N and ^{13}C chemical shift assignments were made from the standard suite of triple resonance experiments as described previously (*9*). NOE-derived distance restraints were obtained from a 2D NOESY, and ϕ angle constraints were calculated on the basis of $^{3}J_{HNH}\alpha$ coupling constants measured from a 3D HNHA (*10*) or using the INFIT module of DYANA (*11*). Additional ϕ and ψ restraints were included on the basis of an analysis of H^N, ^{15}N, $^{13}C\alpha$, and $^{13}C\beta$ chemical shifts in the program TALOS (*12*). All NMR data were processed

using XWINNMR (Bruker, Karlsruhe) and analyzed with the program XEASY (*13*).

Structure Calculations. NOE-derived distance restraints obtained from the 2D ^{1}H-NOESY spectra were calibrated using the CALIBA module of DYANA (*11*). Iterative manual assignment of NOEs and preliminary structure calculations were carried out using DYANA (*11*). Further refinement of the structure was conducted in an automated manner, using ARIA (*14*, *15*) implemented in CNS 1.1 and using the standard protocol provided with the software. The final assignments made by ARIA were checked and corrected manually where necessary.

Calculations were carried out in the simplified all-hydrogen PARALLHDG5.2 force field with nonbonded interactions modeled by the PROLSQ force field (*16*); floating chirality assignment was used for all methylene and isopropyl groups. Finally, the 50 lowest energy structures were refined in a 9-Å shell of water molecules (*17*). The 20 conformers with the lowest value of E_{tot} were visualized and analyzed using the programs MOLMOL (*18*), PROCHECK (*19*), and WhatIf (*20*). The coordinates of these 20 conformers have been deposited in the Protein Data Bank with the accession code 2HI3. Chemical shift data have been submitted to the BioMagResBank with the code 7203.

Gel Retardation Assay Using Pentaprobe. Purified recombinant HOP was prepared as described above. Electrophoretic mobility shift assay (EMSA) reactions were set up in a total volume of 30 μL, comprising approximately 1 pmol of ^{32}P labeled probe, ~100 ng of recombinant protein, 10 mM Hepes at pH 7.8, 50 mM KCl, 5 mM $MgCl_2$, 1 mM EDTA and 5% glycerol, and 25 μg/mL dIdC. After incubation on ice for 10 min, the samples were loaded onto a 6% native polyacrylamide gel made up in 0.5× TBE. The gel was then subjected to electrophoresis at 15 V/cm at 4 °C for 3 h, dried, analyzed, and quantified when necessary using a PhosphoImager (Molecular Dynamics).

Preparation of HOP Mutants. Initially, a range of HOP point mutants were created using a random mutagenesis kit (GeneMorph, Qiagen), and their sequences were confirmed. Subsequently, specific mutants were made using site-directed mutagenesis (Qiagen). HOP mutants were overexpressed and purified in the same manner as the wild-type protein.

Transient Transfection Assays. Transient transfection assays were performed as described (*1*). Briefly, COS7 cells were transfected using FuGene6 (Roche) as directed. The reporter construct used was −5000 SM22α-Luc (kindly provided by M. Parmacek). Flag-myocardin and HA-SRF were kindly provided by E. Olson and by M Parmacek, respectively. PSVβ-galactosidase (Promega) was used to normalize for transfection efficiencies and pcDNA3 was used to equalize the amount of DNA in each transfection. Western blotting of whole cell lysates with an anti-myc antibody was used to determine expression levels for wild-type and mutant proteins.

Coimmunoprecipitation and Western Blot Analysis. Coimmunoprecipitation and Western blot analysis were described previously (*4*). 293T cells were transfected with Flag-tagged wild-type or mutant HOP and HA-tagged SRF. The cells were lysed and pre-cleared, and then 2 ug of anti-FLAG antibody was added. The immunocomplex was recovered by protein A Sepharose beads (Santa Cruz Biotechnology, Inc.). SRF was detected with an anti-HA antibody after separation on SDS−PAGE gels.

Indirect Immunofluorescence. Intracellular localization of the HOP mutant was visualized with fluorescence. COS cells were transfected with wild type or mutant HOP, fixed with 3.7% paraformaldeyde in PBS (pH 7.4), permeabilized with 0.2% Triton-X, washed, and blocked in PBS containing 10% normal goat serum for 10 min. After washing, HOP in the cells was probed by incubating with anti-myc antibody (1:200) overnight at 4 °C. The cells were then washed three times with PBS containing 0.1% Tween (PBS-T) and visualized with Alexa Fluor 568-conjugated goat anti-mouse IgG antibody (Molecular Probes, Eugene, Oregon, USA) diluted in blocking solution (1:500) for 30 min. The cells were washed with PBS-T twice, and the nuclei were then counterstained with DAPI (Molecular Probes) and observed with fluorescent microscopy (Olympus IX50).

RESULTS

HOP Is Conserved in Invertebrates. Although an earlier analysis of DNA sequences found no evidence for HOP-related proteins in organisms other than vertebrates (*1*), an analysis of current DNA sequence databases revealed the existence of homeodomain-only proteins in several invertebrates (Figure 1A), namely, *Apis mellifera* (honeybee), *Locusta migratoria* (migratory locust), *Acyrthosiphon pisum* (pea aphid), *Mesobuthus gibbosus* (Mediterranean checkered scorpion), and *Ixodes scapularis* (black-legged tick). This demonstrates that HOP-type proteins are more ancient than previously thought. An extensive search of current sequence data confirmed the earlier assertion (*1*) that these are the only proteins that consist entirely of a homeodomain. All other homeodomain-containing proteins identified are > ~200 amino acids in length and contain additional sequence elements.

HOP Is a Monomeric Homeodomain in Solution. As part of our efforts to understand the molecular mechanism through which HOP acts to effect the repression of SRF-regulated genes, we overexpressed and purified full-length HOP. A circular dichroism spectrum (not shown) revealed that HOP is predominantly α-helical in solution, and sedimentation equilibrium data show that HOP is monomeric in solution (Figure 2A). We next used multidimensional, multinuclear NMR spectroscopy to determine the solution conformation of HOP. HOP gave rise to good quality NMR data; a ^{15}N-HSQC spectrum is shown in Figure 2B. Figure 3A shows an overlay of the 20 lowest energy structures calculated from ARIA (RMSD = 0.74 Å for backbone C′, C$^\alpha$ and N atoms of residues 13−65), which were used to represent the solution structure of HOP. No constraint violations of greater than 0.5 A for NOEs or 5° for dihedral angles were observed; structural statistics for the ensemble are given in Table 1. As predicted from amino acid sequence analysis, HOP forms a classical homeodomain fold. An overlay of the backbone of HOP (best-fitting helical regions; N, C$^\alpha$, and C atoms; Figure 3C) with the homeodomain from the paired domain of Pax6 has an RMSD of 1.2 Å, confirming that the structure of HOP falls into the homeodomain class of protein domains.

The structure of HOP comprises three α-helices connected by short loops, together with disordered *N*- and *C*-terminal regions. The helices are packed to form a hydrophobic core

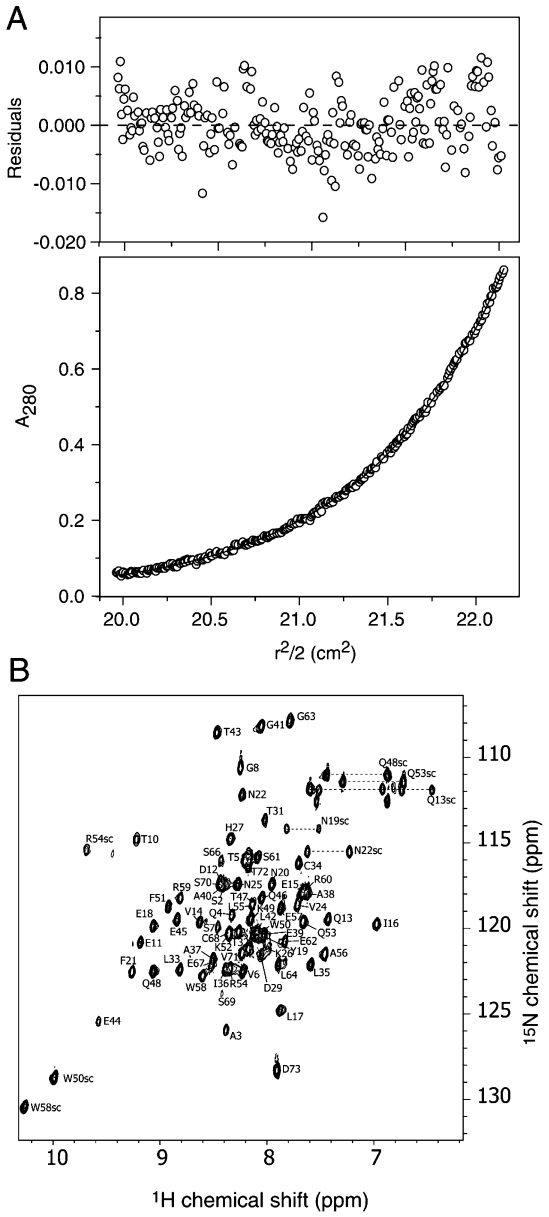

FIGURE 2: Characterization of HOP. (A) Example of data from a sedimentation equilibrium experiment is shown. The bottom panel shows the concentration distribution of HOP (10 μM, 35 000 rpm) fitted to an ideal, single-species model. The top panel shows the residuals to the fit from the same data set. The fits were obtained by simultaneously fitting two datasets recorded at three different speeds (30 000, 35 000, and 42 000 rpm). r is the radial position from the center of the rotor. Experiments were carried out at 25 °C. (B) ^{15}N-HSQC of HOP (0.5 mM, 50 mM phosphate at pH 6.5), recorded at 298 K.

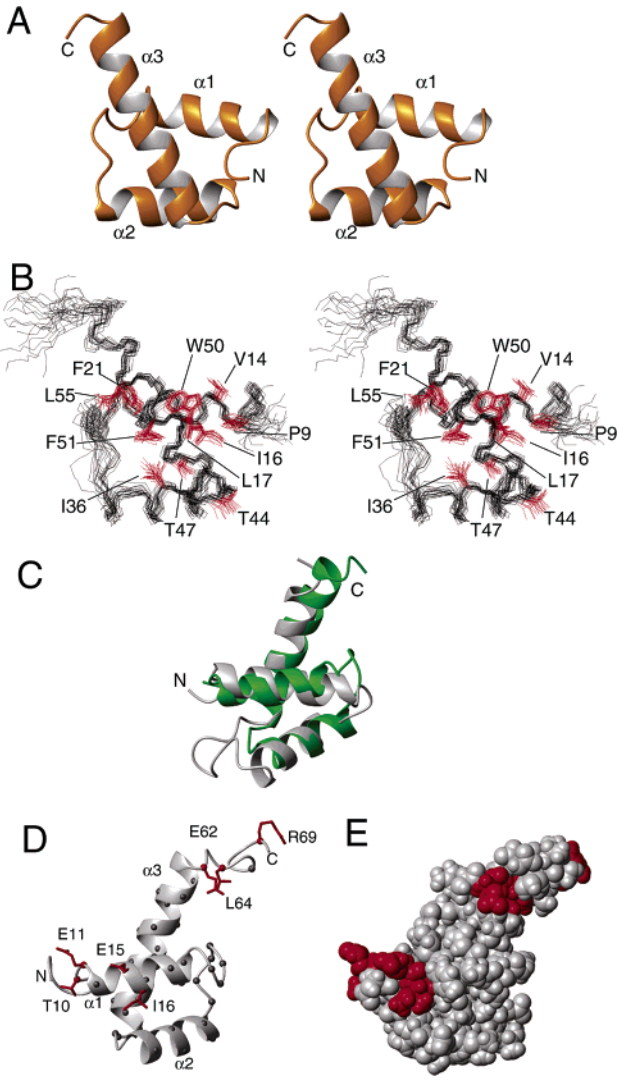

FIGURE 3: Solution structure of HOP. (A) Stereo ribbon diagram of HOP; residues 7−65 are shown. (B) Stereo diagram of the 20 lowest energy structures of HOP. Well-ordered side chains are shown in red and labeled. (C) Overlay of the structures of HOP (green) and the homeodomain from Pax6 (6PAX; gray) over C$^{\alpha}$, C, and N atoms, giving an RMSD of 1.2 Å. (D) Structural mapping of the amino acids required for the function of HOP as a repressor of the SM22α promoter. Residues that when mutated gave rise to significant reductions in repression ability are highlighted in red. The α-carbons of residues that when mutated (either as single point mutations or as part of a multiple-residue mutant) did not yield a significant change in repression ability are shown as gray spheres. (E) Space-filling representation of D.

containing several residues from helix 2, together with L17, F21, W50, and F51; these latter four residues are highly conserved in the homeodomain family. The first two helices are up to 14 and 11 residues in length (depending on which member of the ensemble is examined), whereas the *C*-terminal helix is 18 residues in length. It is this *C*-terminal helix that in all other characterized homeodomains recognizes ATTA (or related) sequences in double-stranded DNA (*21*). However, the absence of the highly conserved asparagine and arginine residues (at positions 53 and 55, respectively, in Figure 1) in HOP as well as other DNA-binding residues such as Y27 and I49 led to the prediction that HOP would not be able to recognize DNA in the same way (*1*). Indeed, site-selection experiments failed to identify a consensus DNA sequence to which HOP could bind. However, because of the difficulties associated with interpreting a negative result in a site-selection assay, we sought additional evidence pertinent to whether HOP could recognize ATTA motifs in double-stranded DNA. We have recently developed Pentaprobe, a series of six ~100-bp oligonucleotides that contain all possible 5-bp sequences (*22*). This set of probes contains eight ATTA sites in different sequence contexts. We carried out gel retardation assays with double-stranded versions of these oligonucleotides. HOP was unable to bind to any of the probes (Supporting Information), confirming that it is unable to function as a typical DNA-binding homeodomain.

Selected Mutations in HOP Abrogate HOP-Mediated Transcriptional Repression. To gain insight into the relation-

Table 1: Structural Statistics for the Family of 20 HOP Structures

Experimental Restraints	
NOEs	783
TALOS restraints	144
dihedral angles (ϕ)	48
Mean RMSD from Idealized Covalent Geometry[a]	
bonds (Å)	0.00328 ± 0.00008
angles (°)	0.43 ± 0.015
impropers (°)	1.18 ± 0.12
Mean RMSD from Experimental Restraints	
NOE (Å)	0.016 ± 0.004
dihedral angles (°)	0.36 ± 0.15
Mean Energies after Water Refinement (kJ mol^{-1})	
E_{NOE}	15 ± 7
E_{cdih}	0.9 ± 0.9
E_{vdW}	-264 ± 11
E_{elec}	-2990 ± 90
E_{bond}	12.8 ± 0.7
$E_{improper}$	33 ± 5
E_{angle}	63 ± 4
E_{total}	-3130 ± 90
Atomic RMS Differences versus the Mean (Å)	
backbone atoms (9−33)	0.64 ± 0.17
all heavy atoms (9−33)	1.18 ± 0.17

[a] The statistics (average ± SD) are calculated for the bundle of the 20 best-energy conformers.

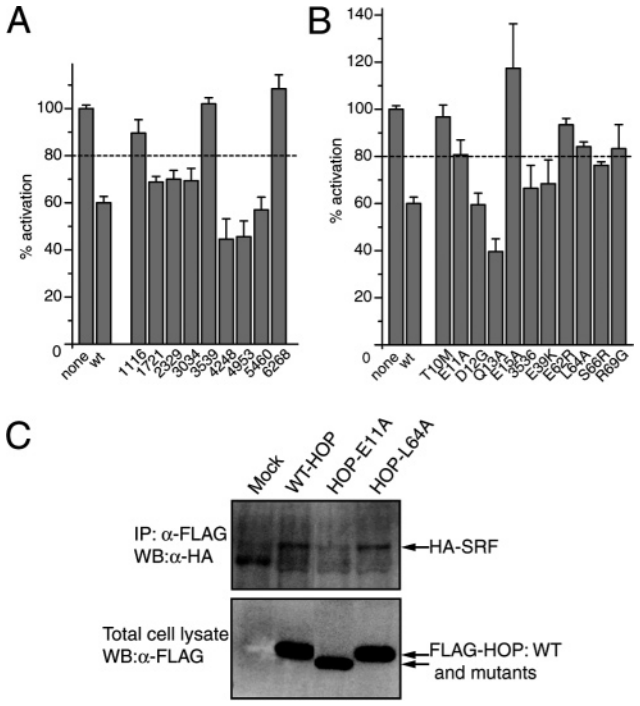

FIGURE 4: Functional analysis of HOP. Mutants of HOP are unable to repress myocardin-induced transactivation of the SM22α promoter in transient transfection assays. Each bar is an average of at least six replicates. Mutations that reduce HOP repression so that activation of the reporter gene is >80% of that obtained without the addition of HOP (above the dotted line) are considered to have the largest effect on HOP activity. (A) Multi-residue mutations, spanning most of the HOP sequence. (B) Point mutants, designed on the basis of the results of A. (C) Co-immunoprecipitation experiment showing that the E11A mutation reduces the ability of HOP to immunoprecipitate SRF. In contrast, L64A does not impair the interaction. Lanes are spliced together from the same gel.

ship between the structure of HOP and its function, we took advantage of the observation that HOP inhibits SRF-dependent transcriptional activation. Myocardin, a recently described SRF transcriptional co-activator (*23*), induces high levels of reporter gene expression when transfected into SRF-containing cells together with a reporter gene driven by an enhancer containing SRF binding sites. We therefore assayed the ability of a panel of multiresidue HOP mutants to repress myocardin-induced activation of such a reporter, an SM22α-luciferase reporter construct that contains adjacent SRF binding sites (Figure 4A). The series of 5−7 residue mutants essentially spanned the entire coding sequence (Figure 1). Wild-type HOP was able to repress myocardin-induced activation of this reporter construct, consistent with previous results (*1*). Surprisingly, the majority of the multiple-residue HOP mutants repressed activation of the reporter construct at wild-type levels, indicating that the HOP fold is rather robust. Three of the multiple-residue mutants, 1116, 3539, and 6268, significantly reduced HOP activity.

Although all three of these mutants express at normal levels in the transient transfection experiments (Supporting Information), the 1116 and 3539 mutants appear to induce some destabilization of the protein structure, judging from NMR data and stability in our standard *E. coli* expression system. Therefore, to probe these and other regions in more detail, a series of point mutants was constructed and their activity in the transfection assay tested (Figure 4B). These data showed that the mutations to residues T10M, E11A, E15A, E62R, L64A, S66R, and R69G reduced the repressive activity of HOP.

Mutations that cause a loss of function may do so through the substitution of a surface residue that is essential for the formation of a specific intermolecular interaction. Alternatively, a mutation may simply disrupt the global fold of the target protein. Thus, we recorded circular dichroism and 1D ^{1}H NMR spectra (Supporting Information) of purified recombinant versions of the functionally impaired point mutants in order to ascertain whether they were correctly folded. All of these mutants listed above were well folded (spectra were not recorded for R69G, given that this residue lies well within a disordered region of the protein), indicating that each of the amino acid substitutions most likely exhibited their effect by interfering with a specific contact between HOP and one of its partners. It should be noted that several other mutants (for example, A40T) reduced the repressive activity of HOP but either were not correctly folded or were expressed at substantially reduced levels in the transfection experiments (judging from a Western blot); these mutants have not been considered further.

In Figure 3D and E, the mutations that were made have been mapped onto the HOP structure. It can be seen that residues 62−68 lie in the unstructured region immediately to the *C*-terminal side of helix 3, whereas the residues in helix 1 lie well away from the *C*-terminal region.

To determine which of the two regions of HOP identified in the mutagenic analysis might be responsible for contacting SRF, we carried out an immunoprecipitation experiment using HOP mutants selected from each of the regions. As shown in Figure 4C, the E11A mutation reduced the ability of HOP to immunoprecipitate SRF, whereas L64A had no effect. This indicates that the *N*-terminal region of HOP is likely to be important for its interaction with SRF.

DISCUSSION

Our structural data demonstrate that HOP forms a homeodomain fold that is indistinguishable overall from sequence neighbors such as Pax6. Experiments carried out with the Pentaprobe reagent however indicate that HOP is not able to directly bind to the expected ATTA homeodomain recognition sites in DNA nor to any sequence of five bp or less. Although it is possible that HOP recognizes a longer consensus sequence that is not present in Pentraprobe or that HOP contacts DNA only when bound to an unidentified partner protein, our findings are consistent with the fact that HOP lacks the conserved asparagine and arginine residues located in the DNA-binding helix of all other homeodomain proteins. HOP is thus a genuine homeodomain that does not bind DNA but rather contacts other proteins to function as a transcriptional co-repressor. This observation emphasizes the fact that a number of classes of proteins that are traditionally considered to be sequence specific DNA-binding domains can exhibit additional or alternative functions. For example, classical zinc fingers are one of the most common of all protein domains and have primarily been associated with DNA-binding activity in transcriptional regulators. However, we have previously shown that such domains can instead act as protein recognition motifs (*24*), even when embedded within a TFIIIA-like cluster of three tandem fingers. We have also shown that single GATA-type zinc fingers can function simultaneously as protein- and DNA-binding domains (*25*), and it is likely that these are not isolated cases. An appreciation of the diversity of function of nuclear protein domains such as these is important if we are to have the ability to predict protein function from sequence and structural data; current views often do not take this observed diversity into account.

Given that HOP does not bind to DNA, we have sought to understand the molecular basis for its ability to repress genes activated by SRF and myocardin. Here, we showed that two motifs are important for the repression activity of HOP, one comprising residues from helix 1 and the other comprising the *C*-terminal disordered tail of HOP. Mutations of these residues do not prevent the formation of well-ordered native-like structures, demonstrating that they are not required for the structural integrity of HOP but rather are likely to represent points of contact for protein partners such as SRF- and HDAC-containing complexes. Results from co-immunoprecipitation experiments indicate that the *N*-terminal motif is important for the assembly of a complex containing both HOP and SRF.

Attempts to more stringently document a direct interaction between HOP and SRF using GST- or MBP-pulldowns and purified recombinant HOP and SRF were not feasible largely because of the poor solution properties of the recombinant SRF MADS box domain (which homoaggregates and has a propensity to nonspecifically bind to other proteins). Furthermore, mammalian two-hybrid experiments were unable to detect an interaction between SRF and HOP, suggesting that unknown additional factors present in transfected cells may facilitate HOP−SRF and/or HOP−HDAC interactions. Our attempts to purify a HOP nuclear complex using large-scale protein purification techniques have thus far been unsuccessful, though the identification of additional HOP-interacting proteins remains a focus of ongoing studies.

It has been shown previously in GST-pulldown experiments that HOP can interact with SRF, when at least one of the proteins is produced from in vitro transcription/translation from mammalian reticulocyte lysates (*1*, *2*). Similarly, epitope tagged HOP and SRF can be co-immunoprecipitated from mammalian cells (*1*, *2*). A simple model for the action of HOP would, therefore, be that it binds to promoter-bound SRF and represses transcriptional activity through the recruitment of HDAC activity. This model, however, does not appear to be consistent with the observation that in vitro translated HOP can reduce the ability of SRF to bind to DNA in a gel shift assay (*1*, *2*).

It is perhaps notable that we have been unable to observe a direct effect of HOP on DNA binding by the SRF MADS domain when both proteins are expressed and purified recombinantly from bacteria (data not shown). That is, the addition of HOP (up to a ~30-fold molar excess) does not result in any significant changes in SRF DNA-binding activity in a gel shift. This result contrasts that obtained using protein expressed from reticulocyte lysates, and it is possible that additional factors are involved in the mechanism of HOP activity. It is also possible that the insolubility of a HOP−SRF complex under the conditions used for the earlier gel shift assays might have resulted in an apparent reduction in DNA binding by SRF. Indeed, we have observed solubility problems for mixtures of HOP and the SRF MADS domain (data not shown). It remains possible that HOP functions to reduce SRF-dependent gene activation by competing with other SRF−homeodomain interactions (*26*) and by recruiting co-repressors, including HDACs.

In summary, we have determined the solution structure of HOP, shown that it forms a homeodomain fold, and delineated surfaces on the protein that are important for its function. Further work will be required to understand the molecular details underlying HOP activity and define its interacting partners.

ACKNOWLEDGMENT

Merlin Crossley is acknowledged for helpful discussions. Bill Bubb is acknowledged for skilful maintenance of the 600-MHz NMR spectrometer at USyd.

SUPPORTING INFORMATION AVAILABLE

Circular dichroism and 1D ^{1}H NMR spectra of purified recombinant versions of the functionally impaired point mutants. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Chen, F., Kook, H., Milewski, R., Gitler, A. D., Lu, M. M., Li, J., Nazarian, R., Schnepp, R., Jen, K., Biben, C., Runke, G., Mackay, J. P., Novotny, J., Schwartz, R. J., Harvey, R. P., Mullins, M. C., and Epstein, J. A. (2002) Hop is an unusual homeobox gene that modulates cardiac development, *Cell 110*, 713−723.
2. Shin, C. H., Liu, Z. P., Passier, R., Zhang, C. L., Wang, D. Z., Harris, T. M., Yamagishi, H., Richardson, J. A., Childs, G., and Olson, E. N. (2002) Modulation of cardiac growth and development by HOP, an unusual homeodomain protein, *Cell 110*, 725−735.
3. Pashmforoush, M., Lu, J. T., Chen, H., Amand, T. S., Kondo, R., Pradervand, S., Evans, S. M., Clark, B., Feramisco, J. R., Giles, W., Ho, S. Y., Benson, D. W., Silberbach, M., Shou, W., and Chien, K. R. (2004) Nkx2-5 pathways and congenital heart

disease; loss of ventricular myocyte lineage specification leads to progressive cardiomyopathy and complete heart block, *Cell 117*, 373−386.
4. Kook, H., Lepore, J. J., Gitler, A. D., Lu, M. M., Wing-Man Yung, W., Mackay, J., Zhou, R., Ferrari, V., Gruber, P., and Epstein, J. A. (2003) Cardiac hypertrophy and histone deacetylase-dependent transcriptional repression mediated by the atypical homeodomain protein Hop, *J. Clin. Invest. 112*, 863−871.
5. Cai, M., Huang, Y., Sakaguchi, K., Clore, G. M., Gronenborn, A. M., and Craigie, R. (1998) An efficient and cost-effective isotope labeling protocol for proteins expressed in *Escherichia coli*, *J. Biomol. NMR 11*, 97−102.
6. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques, *Biophys. J. 36*, 575−588.
7. Hayes, D. B., Laue, T., and Philo, J. (1995) *SEDNTERP*, USA.
8. Liew, C. K., Kowalski, K., Fox, A. H., Newton, A., Sharpe, B. K., Crossley, M., and Mackay, J. P. (2000) Solution structures of two CCHC zinc fingers from the FOG family protein U-shaped that mediate protein−protein interactions, *Structure 8*, 1157−1166.
9. Deane, J., Mackay, J., Kwan, A., Sum, E., Visvader, J., and Matthews, J. (2003) Structural basis for the recognition of ldb1 by the N-terminal LIM domains of LMO2 and LMO4, *EMBO J. 22*, 2224−2233.
10. Vuister, G., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bind J(HN−Ha) coupling constants in 15N-enriched proteins, *J. Am. Chem. Soc. 115*, 7772−7777.
11. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA, *J. Mol. Biol. 273*, 283−298.
12. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
13. Bartels, C., Xia, T.-H., Billeter, P., Guntert, P., and Wuthrich, K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules, *J. Biomol. NMR 5*, 1−10.
14. Nilges, M. (1995) Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE cross-peaks and disulphide connectivities, *J. Mol. Biol. 245*, 645−660.
15. Nilges, M., Macias, M. J., O'Donoghue, S. I., and Oschkinat, H. (1997) Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from beta-spectrin, *J. Mol. Biol. 269*, 408−422.
16. Linge, J. P., and Nilges, M. (1999) Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation, *J. Biomol. NMR 13*, 51−59.
17. Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R., and Klein, M. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.
18. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures, *J. Mol. Graph. 14*, 51−5, 29−32.
19. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
20. Vriend, G., and Sander, C. (1993) Quality control of protein models: directional atomic contact analysis, *J. Appl. Crystallogr. 26*, 47−60.
21. Billeter, M. (1996) Homeodomain-type DNA recognition, *Prog. Biophys. Mol. Biol. 66*, 211−225.
22. Kwan, A. H., Perdomo, J., Czolij, R., Mackay, J. P., and Crossley, M. (2003) Pentaprobe: a comprehensive sequence for the one-step detection of DNA-binding activities, *Nucleic Acids Res. 31*, e124.
23. Wang, D., Chang, P. S., Wang, Z., Sutherland, L., Richardson, J. A., Small, E., Krieg, P. A., and Olson, E. N. (2001) Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor, *Cell 105*, 851−862.
24. Simpson, R. J. Y., Lee, S. H. Y., Bartle, N., E. Y., S., Visvader, J. E., Matthews, J. M., Mackay, J. P., and Crossley, M. (2004) A classical zinc finger from FOG mediates an interaction with the coiled-coil of TACC3., *J. Biol. Chem. 279*, 39789−39797.
25. Liew, C. K., Simpson, R. J. W., Kwan, A. H., Crofts, L., Loughlin, F., Matthews, J., Crossley, M., and Mackay, J. P. (2005) Zinc fingers as protein recognition motifs: structural basis for the GATA-1/Friend of GATA interaction, *Proc. Natl. Acad. Sci. U.S.A. 102*, 583−588.
26. Chen, C. Y., and Schwartz, R. J. (1996) Recruitment of the tinman homolog Nkx-2.5 by serum response factor activates cardiac alpha-actin gene transcription, *Mol. Cell Biol. 16*, 6372−6384.

BI060641S