PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

SHORT COMMUNICATION

The basic helix–loop–helix region of the transcriptional repressor hairy and enhancer of split 1 is preorganized to bind DNA

Matija Popovic,[1] Hans Wienk,[2] Maristella Coglievina,[1] Rolf Boelens,[2] Sándor Pongor,[1] and Alessandro Pintar[1]*

[1] Protein Structure and Bioinformatics Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), AREA Science Park, I-34149 Trieste, Italy
[2] Bijvoet Center for Biomolecular Research, Utrecht University, 3584 CH Utrecht, the Netherlands

ABSTRACT

Hairy and enhancer of split 1, one of the main downstream effectors in Notch signaling, is a transcriptional repressor of the basic helix–loop–helix (bHLH) family. Using nuclear magnetic resonance methods, we have determined the structure and dynamics of a recombinant protein, H1H, which includes an N-terminal segment, b1, containing functionally important phosphorylation sites, the basic region b2, required for binding to DNA, and the HLH domain. We show that a proline residue in the sequence divides the protein in two parts, a flexible and disordered N-terminal region including b1 and a structured, mainly helical region comprising b2 and the HLH domain. Binding of H1H to a double strand DNA oligonucleotide was monitored through the chemical shift perturbation of backbone amide resonances, and showed that the interaction surface involves not only the b2 segment but also several residues in the b1 and HLH regions.

Proteins 2014; 00:000–000.

Key words: NMR; secondary structure; backbone dynamics; chemical shift perturbation; conformational selection.

INTRODUCTION

The basic helix–loop–helix (bHLH) motif (ProSite: PS50888; Pfam: PF00010; SMART: HLH) is made of a stretch of 10–15 mainly basic residues followed by two helical segments with amphipathic character, connected by a loop region of variable length. Dimerization of the ~50 residue helix–loop–helix (HLH) domains orients the basic regions in the same direction, thus favoring their interaction with the negatively charged phosphate backbone of double strand DNA (dsDNA). The bHLH motif is widely used by eukaryota to bind dsDNA,[1] and is found in transcriptional activators as well as in transcriptional repressors, usually in combination with other domains, such as the Orange, PAS and leucine zipper (LZ) domains. With more than 100 different human proteins containing a bHLH domain, variations in the amino acids that form the HLH domain, in combination with the possibility to form heterodimers, either with members of the same family or with members of other families, leads to a very flexible and

Additional Supporting Information may be found in the online version of this article.
Grant sponsor: EU (Bio-NMR, European Commission's FP7) for access to the SONNMRLSF European Large-Scale NMR Facility in Utrecht, the Netherlands; Grant number: 261863.
Matija Popovic's current address is Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom.
Matija Popovic and Hans Wienk contributed equally to this work.
*Correspondence to: Alessandro Pintar, Protein Structure and Bioinformatics Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), AREA Science Park, Padriciano 99, I-34149 Trieste, Italy. E-mail: pintar@icgeb.org
Received 30 August 2013; Revised 19 December 2013; Accepted 6 January 2014
Published online 8 January 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24507

totally tunable dsDNA binding system.[2,3] The early structures of myoblast determination protein 1 (MyoD; PDB: 1MDY),[4] upstream stimulatory factor 1 (USF-1; PDB: 1AN4),[5] Myc-associated factor X (Max; PDB: 1AN2, 1HLO),[6,7] phosphate system positive regulatory protein (PHO4; PDB: 1A0A),[8] and sterol regulatory element binding protein 1A (SREBP-1; PDB: 1AM9)[9] homodimers bound to dsDNA unraveled the mode of binding of bHLH proteins to DNA, whereas the crystal structures of the Myc-Max (PDB: 1NKP) and Mad-Max (PDB: 1NLW)[10] heterodimers bound to DNA provided more details on the heterodimer interface. More recently, the structures of the transcription factor E2-alpha/neurogenic differentiation factor 1 (E47/NeuroD1; PDB: 2QL2),[11] circadian locomotor output cycles kaput/brain and muscle ARNT-like 1 (CLOCK/BMAL1; PDB: 4H10),[12] and microphthalmia-associated transcription factor (MITF; PDB: 4ATI, 4ATK)[13] complexes with DNA were also solved. Although the atomic interactions and thermodynamic principles that stabilize a DNA/bHLH protein complex are now known, the dynamic processes that lead to DNA recognition by bHLH proteins, the discrimination of the cognate sequence on the DNA, and the difference between nonspecific and specific binding are less well understood, and depend on a better comprehension of the solution structure and dynamics of bHLH proteins in the absence of DNA. The little information available on the structure and dynamics of free bHLH domains in solution depicts a quite variable scenario. The neurogenin 1 bHLH domain has been described as a monomeric natively unfolded protein, but detailed heteronuclear nuclear magnetic resonance (NMR) data are not available.[14] NMR studies on the E47 isoform of transcription factor E2α homodimer showed that the basic region is mainly disordered, despite some evidence of nascent helix formation, and that the HLH domain is structured, though highly dynamic.[15] Similar results were obtained also for yeast PHO4, with backbone dynamics showing a progressive increase in the rigidity of the basic region, from the N terminus to the start of the first helix.[16] A covalent dimer of MyoD also showed a disordered basic region and a structured HLH domain.[17] Taking into consideration also bHLH-LZ proteins, yet another picture was obtained for viral Myc transforming protein (v-Myc) in solution. In v-Myc, no stable dimer could be observed, the HLH domain is mainly disordered, the basic region displays some residual helical structure, and only the LZ domain is well ordered, despite the overall high backbone flexibility.[18] The obligate heterodimeric partner of Myc, Max, forms instead a well structured homodimer (PDB: 1R05), with the HLH-LZ domains showing little differences with the crystal structures available.[19] In Max, the basic region is mainly disordered, with the exception of three residues close to the start of the HLH domain.

The domain organization of hairy and enhancer of split 1 (HES-1), which includes a bHLH domain, an Orange domain and a C-terminal half has been investigated,[20] but no structure is available for HES-1, either free or bound to DNA, nor for any transcriptional repressor with the same domain architecture. Furthermore, the basic segment of HES-1 is characterized by a proline residue that is conserved in the members of this family (human HES-1, −2, −4, −5, −6, and −7) but is not found in other bHLH proteins. To investigate the structure and dynamics of the bHLH region, we expressed a recombinant protein, H1H, corresponding to residues 27–95 of human HES-1 [Fig. 1(A)], and including an N-terminal region, b1, which contains serine phosphorylation sites reported to be functionally important,[21,22] the basic segment b2, conserved in all bHLH proteins, and the HLH domain. We report here the secondary structure and backbone dynamics of H1H, as determined by NMR, and a structural model of the free H1H dimer, generated through a combination of chemical shift-driven *ab initio* structure generation, restrained torsion angle dynamics and energy minimization. We also report the interaction of H1H with a synthetic dsDNA oligonucleotide, mapped through backbone HN chemical shift perturbation (CSP), and compare the structural and binding properties of HES-1 bHLH domain with other bHLH proteins.

MATERIALS AND METHODS

Sample preparation

The recombinant H1H protein (70 amino acids), corresponding to residues 27–95 of human HES-1 preceded by a start methionine and including the b1 and b2 basic regions and the HLH domain was expressed in *E. coli* BL21(DE3)pLysS cells from a codon-optimized synthetic gene, purified by ion exchange chromatography followed by reverse-phase high-pressure liquid chromatography, and freeze-dried, as described.[20] Uniform ^{15}N- or ^{15}N, ^{13}C-labeling was achieved growing cells in a minimal medium containing 6 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 0.12 g/L $MgSO_4$, 0.01 g/L $CaCl_2$, and 1.7 g/L yeast nitrogen base without amino acids and ammonium sulfate (Difco), 100 μg/mL ampicillin, and 25 μg/mL chloramphenicol supplemented with 0.5 g/L $^{15}NH_4Cl$, 5 g/L D-glucose or 0.5 g/L $^{15}NH_4Cl$, 5 g/L U-$^{13}C_6$ D-glucose for the singly and doubly labelled proteins, respectively. The purified proteins were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis and liquid chromatography-mass spectrometry to assess their purity and correct molecular weight. NMR samples were prepared dissolving the freeze-dried powder in 50 m*M* sodium phosphate buffer, pH 6.5, containing 10% (v/v) D_2O, for a final protein concentration of ∼2 m*M*.

The hASH1c* dsDNA oligonucleotide used for NMR binding studies derives from the promoter region of the transcription factor achaete-scute homolog 1 (ASH1)[23] and binds H1H with high affinity, as measured by electrophoretic mobility shift assay (data not shown and

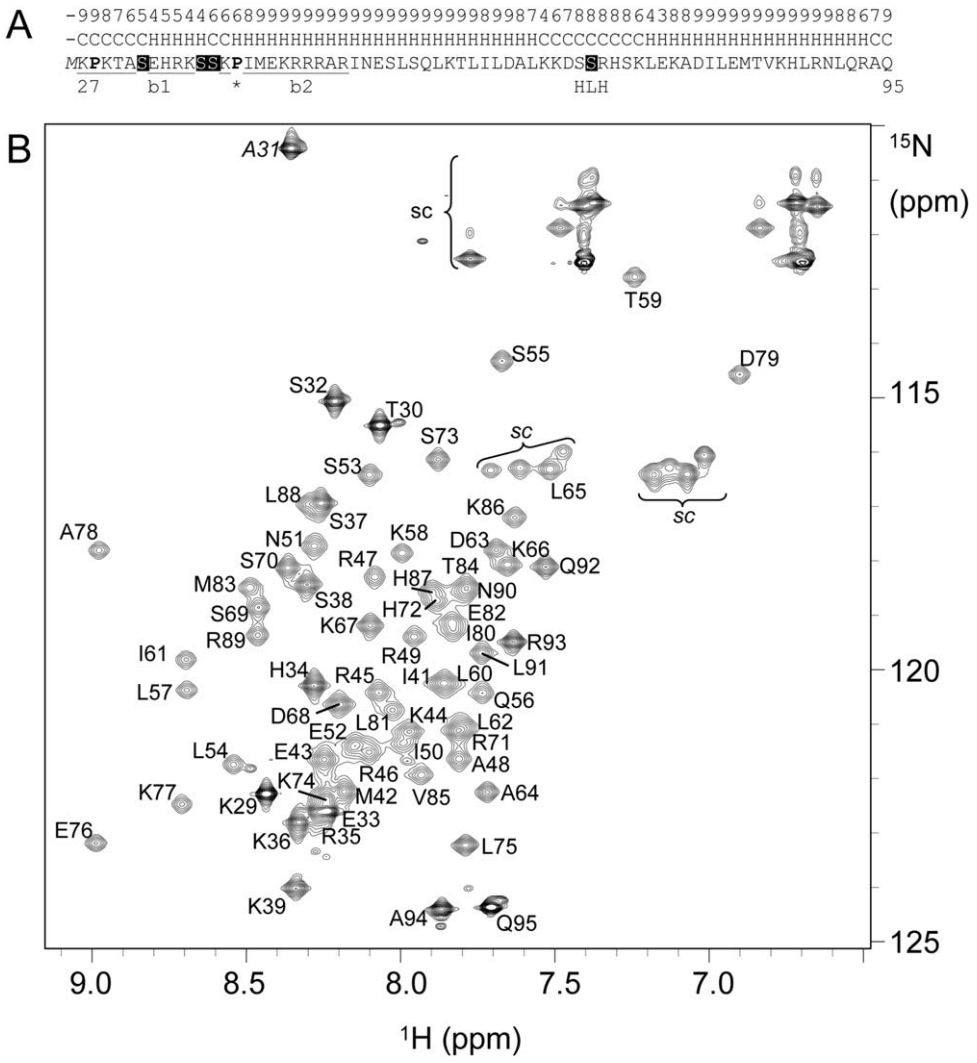

Figure 1

Sequence and $^{1}H–^{15}N$ HSQC NMR spectrum of H1H. (**A**) amino acid sequence, predicted secondary structure (H, helix; C, coil) and prediction confidence of H1H; the b1, b2 and HLH regions are indicated; HES-1 residues found to be phosphorylated are boxed in black; the conserved proline is marked by an asterisk. (**B**) $^{1}H–^{15}N$ HSQC NMR spectrum of H1H recorded at 288 K; assigned residues are labeled; folded peaks in italics; sc, side chains.

Ref. 20). Of the two HES-1 binding sites in this promoter, the first (CACGCAC) can be defined as an atypical class C site as it does not correspond to a canonical N- or E-box, the second binding site in the same promoter is a class C site (CACGCG). The hASH1c* oligonucleotide is a mutant where the class C site was changed into a E-box, class B (CACGTG), and thus contains a palindromic binding motif flanked by nonpalindromic sequences.[24] The 16-mer dsDNA oligonucleotide hASH1c* was obtained by annealing of the corresponding forward (5′-TCCGGCACGTGCCAGG-3′) and reverse (5′-CCTGGCACGTGCCGGA-3′) ssDNA synthetic oligonucleotides (Eurofins MWG Operon, Ebersberg, Germany; high-performance liquid chromatography grade) dissolved in 50 m*M* sodium phosphate, pH 7.4 at equimolar concentrations (5.7 m*M*). The forward and reverse oligonucleotides were annealed by heating at 90°C for 10 min followed by slow cooling to room temperature. The mixture was analyzed on a nondenaturing 15% polyacrylamide (19:1, w/v, acrylamide/bisacrylamide) minigel (Bio-Rad), in 1× 100 m*M* Tris, 90 m*M* boric acid, and 1 m*M* ethylenediaminetetraacetic acid, pH 8.4.

NMR

All NMR experiments on free H1H were carried out at 288 K. Preliminary experiments were acquired on a Bruker Avance 500 MHz spectrometer using ^{15}N-labeled protein. For assignment purposes, standard HNCACB, HNCA, CBCA(CO)NH, HN(CA)CO, HNCO, H(CC)CONH,

C(CC)CONH, HBHA(CO)NH, hCCH-TOCSY, and HcCH-TOCSY experiments[25] were acquired on Bruker Avance 600 and 900 MHz spectrometers equipped with cryogenic probeheads optimized for ^{1}H detection. $^{3}J_{HNH\alpha}$ coupling constants of backbone amides were estimated from HNHA.[26] Data were processed using nmrPipe and analyzed with the CCPN program ANALYSIS (http://www.ccpn.ac.uk). Chemical shifts were referenced to DSS. Deviations of C^{α}, C', H^{α}, and C^{β} chemical shifts from random coil values[27] were used to determine secondary structure regions. T_1 and T_2 relaxation times were estimated from fitting a two-parameter exponential decay to experimental intensities of HN cross peaks acquired with delays of 8, 20, 40, 80, 150, 320, 600, 1000, 1600, 2000, and 2500 ms for T_1 and 0, 0, 17, 17, 34, 50, 67, 118, 150, and 185 ms for T_2 experiments. ^{1}H–^{15}N steady-state heteronuclear NOEs were obtained from the ratio of peak heights in paired spectra collected with and without proton saturation during the relaxation delay. $^{3}J_{HNH\alpha}$ coupling constants were calculated from the intensity ratio between the Hα cross peak and the diagonal HN peak using the formula $I_{H\alpha}/I_{HN} = -\tan^2(2\pi J\xi)$ where J is the $^{3}J_{HNH\alpha}$ coupling constant and $\xi = 13.05$ ms.[26] No correction for relaxation effects was applied.

Titration of H1H with the dsDNA hASH1c* oligonucleotide was carried out at 300 K using H1H dimer/dsDNA molar ratios of 0, 0.1, 0.3, 0.6, and 1.2 in the presence of 0.1 *M* NaCl to improve the solubility of the complex. Chemical shift changes in the HN resonances were monitored through ^{1}H, ^{15}N HSQC spectra and the backbone resonances of H1H in the complex were reassigned using triple resonance experiments. Combined CSP of backbone amides was calculated according to the formula: $CSP = \mathrm{sqrt}[(\Delta\delta_H)^2 + \alpha_N(\Delta\delta_N)^2]$, using a $\Delta\delta_N$ scaling factor (α_N) of 0.102.[28]

The predicted consensus secondary structure and prediction confidence was obtained from the Phyre web server.[29]

Structure calculation

From ^{1}H, ^{15}N, and ^{13}C chemical shift assignments, backbone ϕ and ψ angles were predicted with TALOS+[30] and a model of monomeric H1H (residues M26-Q95) was generated using CS-ROSETTA.[31] The best model was duplicated, and the two monomers were oriented to match the homodimer of the related PHO4 crystal structure (PDB: 1A0A).[8] Together with the TALOS+ results, hydrogen bond restraints, and partially assigned 2D NOESY and [^{1}H;^{15}N]-HSQC-NOESY, the homodimer model was used as input for CYANA 3.0.[32] After eight iterations of automatic NOE assignments and structure calculations, the resulting NMR structures were refined in explicit water with CNS[33] according to the RECOORD protocol.[34] The final ensemble was analyzed using iCing[35] (Supporting Information Table S1). NMR chemical shift assignments were deposited in the Biological Magnetic Resonance Bank database (accession number: 19614). Coordinates for the 20 lowest energy structures were deposited in the Protein Data Bank (PDB code: 2MH3).

RESULTS

The construct used in this study [Fig. 1(A)] includes not only the canonical basic region required for DNA binding (residues 41–49, b2) and the HLH domain but also a basic N-terminal segment (residues 27–39, b1) that contains functionally important phosphorylation sites (S32, S37, and S38). The b1 and b2 regions are separated by P40. The proline residue is conserved in all other members of the HES family (HES-1, −2, −4,-5, −6, and −7), with the exception of HES-3, which lacks the b1 region, and is also found in *Drosophila* deadpan, the orthologue of HES-1.

Preliminary ^{1}H–^{15}N HSQC experiments on H1H (Supporting Information Fig. S1) showed that workable NMR spectra could be obtained only in a narrow range of temperatures around 288 K, with extensive line broadening at lower temperatures, and progressive decrease in the HN chemical shift dispersion at temperatures higher than 298 K, which is consistent with the marginal thermodynamic stability of the H1H dimer and the denaturation midpoint of 37°C, as previously measured by CD.[20] Initial attempts to assign the backbone resonances basing on ^{15}N-edited 3D-TOCSY and NOESY experiments failed because of the relatively low T_2 values, in line with those observed for E47[15]and Max,[19] which leads to poor magnetization transfer along the side chains, and the limited chemical shift dispersion of the H^{α} and H^{β} resonances, consistent with a mainly helical or disordered structure. Backbone assignments and secondary structure were obtained from triple resonance experiments and chemical shift index values, respectively. Despite the good chemical shift dispersion of HN signals [6.9–9.0 ppm; Fig. 1(B)], sequential assignments were complicated by the degeneracy of the amino acid composition. There are no Cys, Gly, Phe, Trp, or Tyr residues in the sequence, and 51% of the residues are accounted for by only four amino acid types, Arg (8), Lys (11), Leu (9), and Ser (8). Deviations from random coil chemical shift values (Fig. 2) suggest that H1H is made of a mainly disordered b1 stretch (K27-K39), a long helical segment (P40-A64) covering the basic region (b2) and the first helix of the HLH domain, a loop (L65-E76) and a second helix (K77-Q92) that displays some fraying at its C terminus. Possibly, a third short one-turn helix (D68-R71) can be identified within the loop region. Coupling constants ($^{3}J_{HNH\alpha}$) calculated, when possible, from an HNHA experiment, are in substantial agreement with this view (Fig. 2).

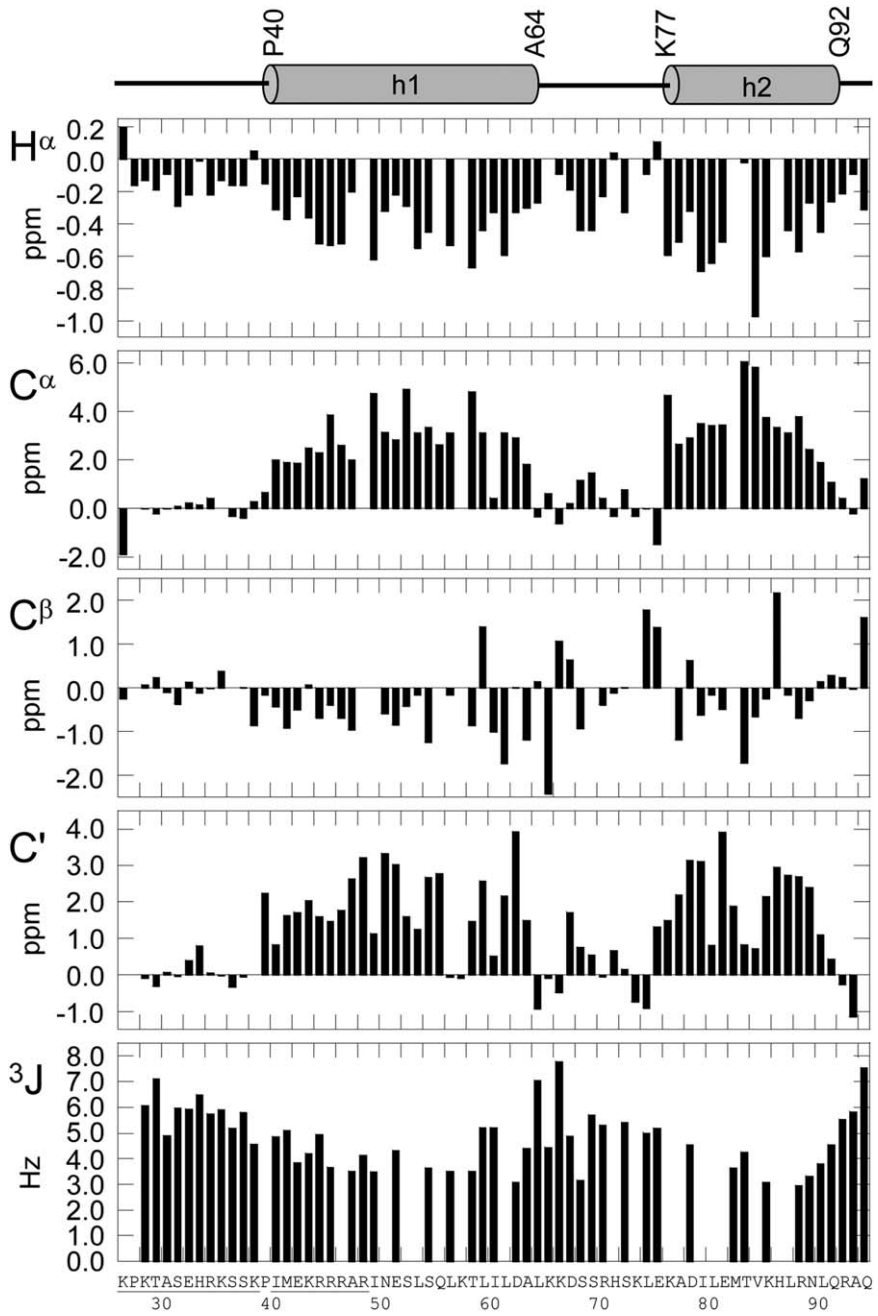

Figure 2

Secondary structure. Secondary structure diagram, chemical shift deviations from random coil values for H^{α}, C^{α}, C^{β}, and C' resonances, and $^{3}J_{HNH\alpha}$ coupling constants.

Based on experimental restraints alone, a dimer structure could not be obtained. Therefore, a structural model of the H1H dimer was calculated from a combination of chemical shift-driven *ab initio* structure generation, restrained torsion angle dynamics and energy minimization. Shortly, a model for the monomer was generated from the assigned chemical shift using CS-ROSETTA. This model showed a disordered N-terminal region followed by a structured core with two major helices (residues P40-A64 and K77-Q92). The loop region also showed a single helical turn spanning residues S69-H72. Again, from NOE restraints it was not possible to define the H1H dimer interface precisely. This problem was overcome observing that the related PHO4 protein shares with H1H a significant sequence similarity in the HLH region and a very similar monomer fold. The H1H

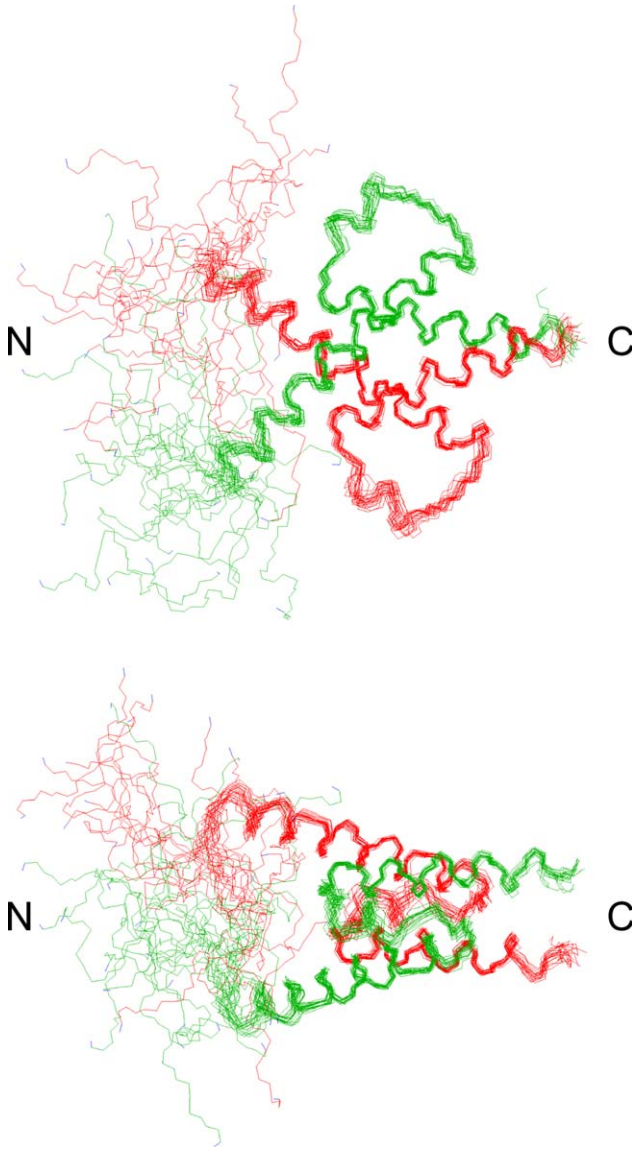

Figure 3

Structure of H1H. Ensemble of the 20 lowest energy models of the H1H dimer in two orthogonal orientations related by a 90° rotation around the *x*-axis; the two chains are colored differently.

monomers were thus oriented as in the PHO4 dimer and this dimer was used as template for automatic NOE assignment and structure calculation in CYANA. With this procedure, 146 intermonomer NOEs were eventually assigned, of which several were unambiguous. After water refinement, an ensemble of 20 models was obtained with no systematic distance (>0.5 Å) and angle (>5°) violations and a Ramachandran distribution (for ordered residues 40–93) of 93% and 7% in the most favored and additionally allowed regions, respectively. The calculated structure confirmed the presence of a helical HLH core preceded by a disordered N-terminal region (Fig. 3). Calculation of the charge potential at the surface showed a highly positive patch formed by b2 residues K44, R45, R46, R47, R49, and K77 of each monomer.

Backbone dynamics, as measured by T_1, T_2 ^{15}N relaxation times and ^{1}H–^{15}N heteronuclear NOEs (Fig. 4), showed that P40 divides H1H in two parts. The first includes the b1 region, which is very flexible and characterized by negative or slightly positive (<0.25) NOE values, with a progressive increase from the N terminus to K39, and a steady decrease in T_2 values in the same region. The second part includes the b2 and HLH subdomains, and is relatively rigid, as judged by systematically higher NOE and lower T_2 values. Within this part, no clear discontinuity can be identified between the b2 and the HLH region. Two relatively rigid segments can be identified, b2/helix1 and helix2, and two more flexible regions, the first corresponding to the loop connecting the two helices and the second corresponding to the C terminus. Overall, there is a good correspondence between backbone dynamics and structure, with flexible regions being mainly disordered, and more rigid regions showing well-defined helical structure and packing.

Binding of H1H to hASH1c* DNA was monitored by the combined CSP of the HN resonances upon titration with a 16-mer dsDNA oligonucleotide, and allowed for the identification of an extensive set of residues involved in binding (Fig. 5). Although several resonances, especially in the b2 region, could not be assigned due to line broadening, this is also the region that is most affected by the interaction with DNA. Several other residues that

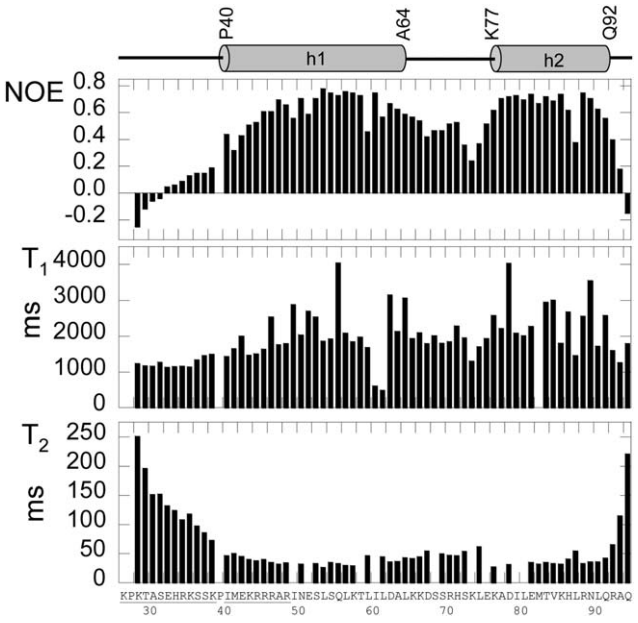

Figure 4

Backbone dynamics. ^{1}H–^{15}N NOE (I/I_{ref}), T_1, and T_2 values calculated for H1H.

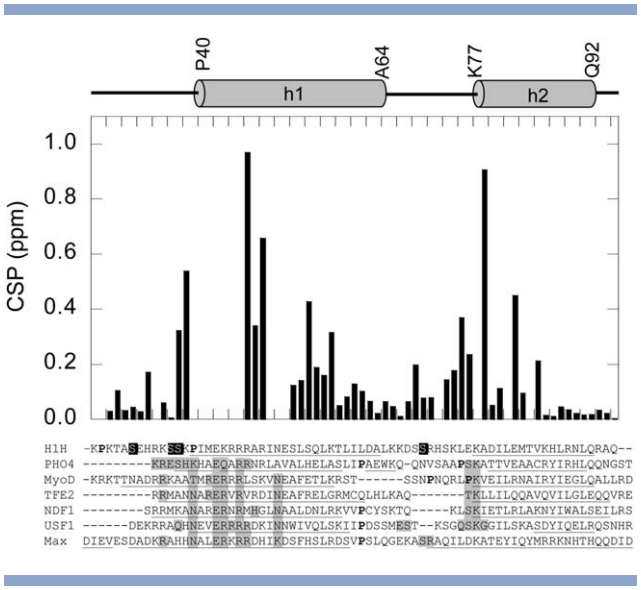

Figure 5

DNA binding. Combined chemical shift perturbation (CSP) values of H1H backbone amides observed upon binding to the hASH1c* dsDNA oligonucleotide plotted vs. the amino acid sequence. The amino acid sequence of other representative bHLH proteins as well as the residues (shaded in grey) interacting with DNA, as derived from the crystal structures of the complexes (PHO4:1a0a; MyoD:1mdy; TFE2/NDF1:2ql2; USF1:1an4; Max:1hlo), are also shown; helical regions are underlined, prolines in bold, HES-1 phosphorylation sites are boxed in black.

are involved were, however, found in the b1 region just before P40 (S38 and K39), in helix1 and also in helix2. Regions that are least affected include the N-terminal part of H1H, the C-terminal part of helix1 and the C-terminal part of helix2.

DISCUSSION

This is the first report of solution NMR studies on the bHLH region of HES-1 and, in general, of a transcription factor with a bHLH-Orange domain architecture. Chemical shifts, coupling constants, and backbone dynamics suggest that, apart from the b1 segment, which is flexible and disordered, the bHLH part of H1H, including the basic stretch of residues (b2) is relatively rigid and structured. The structural model obtained for the free H1H dimer confirms this view. It can be argued that the structure of the H1H dimer is biased towards the structure of PHO4, used to define the initial orientation of the two chains, but the two proteins show a significant degree of amino acid similarity and the experimental constraints found are consistent with the calculated model. Furthermore, the structure of the H1H monomer is fully supported by experimental data. The result obtained for H1H are somewhat different from those reported for other bHLH domains, where the basic region is mainly disordered in the absence of bound DNA, although in some cases NOEs suggest the presence of residual nascent helical turns. It is also in partial contradiction with our limited proteolysis experiments carried out on H1H, which suggested that both the b1 and b2 regions are mainly disordered.[20] This discrepancy can be however explained by the different H1H concentration (μM vs. mM) and temperature (37°C vs. 15°C) used in proteolysis and NMR experiments, respectively. It is known that H1H exists in equilibrium between an unfolded monomer and a structured dimer,[20] and the limited proteolysis conditions likely shift the equilibrium towards the unfolded monomer. It remains to be determined how the bHLH region behaves in the presence of the tandem Orange domain.

It can be speculated that the structural properties of H1H are directed by the presence of the proline at position 40. Because proline residues lack the HN amide proton that normally forms a hydrogen bond with the (i-4) carbonyl and their dihedral angles are conformationally restricted, they are frequently found at the N-cap end of α helices where they act as helix starters. The presence of a proline within the basic segment is a unique feature of transcription factors of the HES family. Prolines have been found in other bHLH proteins but they are usually placed in the loop connecting the two helices, where they might tune the start position and the length of helix2.

The length of helix1 and helix2 of the HLH domain, as well as the length and structure of the loop connecting the two helices vary among bHLH proteins, as seen from their complexes with dsDNA (Supporting Information Fig. S2). While the length and relative orientation of helix1 in the dimer is however restricted by the geometry of bound dsDNA, the structure of the loop and the length and orientation of helix2 are more variable and define the dimer interface. Overall, the topology of H1H, in terms of limits and length of the helical segments, is more similar to that determined for members of the bHLH group than to the topology of bHLH-LZ proteins, where helix1 can be longer and helix2 usually shows no discontinuity with the adjacent helical LZ domain (Supporting Information Fig. S2).

A common topology and mode of binding to dsDNA was also confirmed by the identification of DNA-interacting residues. Although not all the residues belonging to the H1H/DNA interface could be derived unambiguously, a common pattern could be identified, involving several residues in b1, b2, helix1 and the beginning of helix2, similar to other bHLH proteins for which the crystal structure of the complex has been determined. Of the residues displaying a significant (>0.2 ppm) CSP value upon dsDNA binding, E82 and V85 are far from the binding interface, which suggests that these residues could be involved in a change in the dimer interface occurring upon dsDNA binding. Extensive line broadening observed for some HN resonances, possibly arising from conformational exchange, suggests that some key interactions might

undergo intermediate exchange. Binding studies with different oligonucleotides or carried out in different experimental conditions will help in elucidating this point.

Mapping of the H1H/DNA interface also allows for speculations about the functional role of HES-1 phosphorylation. Transcriptional repression or activation mediated by bHLH proteins depends heavily on post-translational modifications that may work either directly, affecting the affinity for DNA, or indirectly, modulating the interaction with other proteins that are recruited to form a transcriptional activator/repressor complex. Different types of post-translational modifications have been found in bHLH proteins, such as lysine acetylation, phosphorylation-dependent acetylation, O-GlcNAc modification, proline hydroxylation, and Ser/Thr phosphorylation.[36] In HES-1, several functionally important phosphorylation sites (S32, S37, S38, S70, and S126) have been identified. S32 in b1, as well as S70 in the HLH domain and S126 in the Orange domain fit the K/RXXS/T phosphorylation motif targeted by calcium/calmodulin-dependent protein kinase type II delta subunit (CaMKIIδ) and were found to be important for the conversion of HES-1 from a transcriptional repressor to a transcriptional activator for neuronal differentiation.[22] S37 and S38 were found to be phosphorylated *in vitro* and in PC12 cells by protein kinase C, with the consequent abolition of DNA binding.[23] The targets of CaMKIIδ phosphorylation, S32 and S70, do not seem to be directly involved in DNA binding, as judged from their CSP values, and it is thus more probable that they are involved in protein–protein interactions. Whereas S32 belongs to a disordered region, S70 is located in the putative short helix in the loop region, and phosphorylation at this site may be associated with local conformational changes, domain rearrangements, and interactions with other proteins. On the contrary, S37 and especially S38 are located in a patch, including K39, that is directly involved in DNA binding as measured by CSP values and as suggested from limited proteolysis experiments.[20] Phosphorylation of these residues adds negative charges that could lead to unfavorable interactions with the negatively charged phosphate backbone of DNA, thus abolishing DNA binding.

In conclusion, we used NMR spectroscopy to investigate the structure and dynamics of a recombinant protein (H1H) that contains the DNA binding region of the transcriptional repressor HES-1, and identified H1H residues that are involved in the binding of the protein to dsDNA. The characterization of the structure and dynamics of H1H forms a good basis for future studies on DNA recognition by HES-1.

ACKNOWLEDGMENT

We thank Nicola D'Amelio (Cluster in Biomedicine, Basovizza, Trieste, Italy) for recording preliminary NMR spectra.

REFERENCES

1. Skinner MK, Rawls A, Wilson-Rawls J, Roalson EH. Basic helix–loop–helix transcription factor gene family phylogenetics and nomenclature. Differentiation 2010;80(1):1–8.
2. Amoutzias GD, Robertson DL, Van de Peer Y, Oliver SG. Choose your partners: dimerization in eukaryotic transcription factors. Trends Biochem Sci 2008;33(5):220–229.
3. Fischer A, Gessler M. Delta-Notch—and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors. Nucleic Acids Res 2007;35(14):4583–4596.
4. Ma PC, Rould MA, Weintraub H, Pabo CO. Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. Cell 1994;77(3):451–459.
5. Ferre-D'Amare AR, Pognonec P, Roeder RG, Burley SK. Structure and function of the b/HLH/Z domain of USF. EMBO J 1994;13(1): 180–189.
6. Ferre-D'Amare AR, Prendergast GC, Ziff EB, Burley SK. Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. Nature 1993;363(6424):38–45.
7. Brownlie P, Ceska T, Lamers M, Romier C, Stier G, Teo H, Suck D. The crystal structure of an intact human Max–DNA complex: new insights into mechanisms of transcriptional control. Structure 1997; 5(4):509–520.
8. Shimizu T, Toumoto A, Ihara K, Shimizu M, Kyogoku Y, Ogawa N, Oshima Y, Hakoshima T. Crystal structure of PHO4 bHLH domain–DNA complex: flanking base recognition. EMBO J 1997; 16(15):4689–4697.
9. Parraga A, Bellsolell L, Ferre-D'Amare AR, Burley SK. Co-crystal structure of sterol regulatory element binding protein 1a at 2.3 Å resolution. Structure 1998;6(5):661–672.
10. Nair SK, Burley SK. X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. Cell 2003;112(2):193–205.
11. Longo A, Guanga GP, Rose RB. Crystal structure of E47-NeuroD1/beta2 bHLH domain-DNA complex: heterodimer selectivity and DNA recognition. Biochemistry 2008;47(1):218–229.
12. Wang Z, Wu Y, Li L, Su XD. Intermolecular recognition revealed by the complex structure of human CLOCK-BMAL1 basic helix–loop–helix domains with E-box DNA. Cell Res 2013;23(2):213–224.
13. Pogenberg V, Ogmundsdottir MH, Bergsteinsdottir K, Schepsky A, Phung B, Deineko V, Milewski M, Steingrimsson E, Wilmanns M. Restricted leucine zipper dimerization and specificity of DNA recognition of the melanocyte master regulator MITF. Genes Dev 2012; 26(23):2647–2658.
14. Aguado-Llera D, Goormaghtigh E, de Geest N, Quan XJ, Prieto A, Hassan BA, Gomez J, Neira JL. The basic helix–loop–helix region of human neurogenin 1 is a monomeric natively unfolded protein which forms a "fuzzy" complex upon DNA binding. Biochemistry 2010;49(8):1577–1589.
15. Fairman R, Beran-Steed RK, Handel TM. Heteronuclear (^{1}H, ^{13}C, ^{15}N) NMR assignments and secondary structure of the basic region-helix–loop–helix domain of E47. Protein Sci 1997;6(1):175–184.
16. Cave JW, Kremer W, Wemmer DE. Backbone dynamics of sequence specific recognition and binding by the yeast Pho4 bHLH domain probed by NMR. Protein Sci 2000;9(12):2354–2365.
17. Starovasnik MA, Blackwell TK, Laue TM, Weintraub H, Klevit RE. Folding topology of the disulfide-bonded dimeric DNA-binding domain of the myogenic determination factor MyoD. Biochemistry 1992;31(41):9891–9903.
18. Fieber W, Schneider ML, Matt T, Krautler B, Konrat R, Bister K. Structure, function, and dynamics of the dimerization and DNA-binding domain of oncogenic transcription factor v-Myc. J Mol Biol 2001;307(5):1395–1410.
19. Sauve S, Tremblay L, Lavigne P. The NMR solution structure of a mutant of the Max b/HLH/LZ free of DNA: insights into the

specific and reversible DNA binding mechanism of dimeric transcription factors. J Mol Biol 2004;342(3):813–832.
20. Coglievina M, Guarnaccia C, Pintar A, Pongor S. Different degrees of structural order in distinct regions of the transcriptional repressor HES-1. Biochim Biophys Acta 2010;1804(12):2153–2161.
21. Strom A, Castella P, Rockwood J, Wagner J, Caudy M. Mediation of NGF signaling by post-translational inhibition of HES-1, a basic helix–loop–helix repressor of neuronal differentiation. Genes Dev 1997;11(23):3168–3181.
22. Ju BG, Solum D, Song EJ, Lee KJ, Rose DW, Glass CK, Rosenfeld MG. Activating the PARP-1 sensor component of the Groucho/TLE1 corepressor complex mediates a CaMKinase IIdelta-dependent neurogenic gene activation pathway. Cell 2004;119(6):815–829.
23. Chen H, Thiagalingam A, Chopra H, Borges MW, Feder JN, Nelkin BD, Baylin SB, Ball DW. Conservation of the Drosophila lateral inhibition pathway in human lung cancer: a hairy-related protein (HES-1) directly represses achaete-scute homolog-1 expression. Proc Natl Acad Sci USA 1997;94(10):5355–5360.
24. Iso T, Kedes L, Hamamori Y. HES and HERP families: multiple effectors of the Notch signaling pathway. J Cell Physiol 2003;194(3):237–255.
25. Cavanagh J, Fairbrother WJ, Palmer AGI, Rance M, Skelton NJ. Protein NMR spectroscopy. San Diego, CA: Elsevier Academic Press; 2007.
26. Vuister GW, Bax A. Quantitative J correlation: a new approach for measuring homonuclear three-bond $J_{HNH\alpha}$ coupling constants in ^{15}N-enriched proteins. J Am Chem Soc 1993;115:7772–7777.
27. Wang Y, Jardetzky O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. Protein Sci 2002;11(4):852–861.
28. Schumann FH, Riepl H, Maurer T, Gronwald W, Neidig KP, Kalbitzer HR. Combined chemical shift changes and amino acid specific chemical shift mapping of protein–protein interactions. J Biomol NMR 2007;39(4):275–289.
29. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. Nat Protoc 2009;4(3):363–371.
30. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 2009;44(4):213–223.
31. Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D, Bax A. Consistent blind protein structure generation from NMR chemical shift data. Proc Natl Acad Sci USA 2008;105(12):4685–4690.
32. Guntert P. Automated NMR structure calculation with CYANA. Methods Mol Biol 2004;278:353–378.
33. Brunger AT. Version 1.2 of the crystallography and NMR system. Nat Protoc 2007;2(11):2728–2733.
34. Nederveen AJ, Doreleijers JF, Vranken W, Miller Z, Spronk CA, Nabuurs SB, Guntert P, Livny M, Markley JL, Nilges M, Ulrich EL, Kaptein R, Bonvin AM. RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins 2005;59(4):662–672.
35. Doreleijers JF, Sousa da Silva AW, Krieger E, Nabuurs SB, Spronk CA, Stevens TJ, Vranken WF, Vriend G, Vuister GW. CING: an integrated residue-based structure validation program suite. J Biomol NMR 2012;54(3):267–283.
36. Benayoun BA, Veitia RA. A post-translational modification code for transcription factors: sorting through a sea of signals. Trends Cell Biol 2009;19(5):189–197.