Biochimica et Biophysica Acta xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Different degrees of structural order in distinct regions of the transcriptional repressor HES-1

Maristella Coglievina, Corrado Guarnaccia, Alessandro Pintar *, Sándor Pongor *

Protein Structure and Bioinformatics Group, International Centre for Genetic Engineering and Biotechnology (ICGEB), AREA Science Park, Padriciano 99, I-34149 Trieste, Italy

ARTICLE INFO

Article history:
Received 19 April 2010
Received in revised form 24 August 2010
Accepted 25 August 2010
Available online xxxx

Keywords:
bHLH domain
Orange domain
Intrinsic disorder
Folding upon binding
Circular dichroism
Electrophoretic mobility shift assay

ABSTRACT

HES-1 is a transcriptional repressor of the basic helix-loop-helix (bHLH) family and one of the main downstream effectors in Notch signaling. Its domain architecture is composed of a bHLH region, an Orange domain, and a poorly characterized C-terminal half. We show that different degrees of structural order are present in the different regions of HES-1. The isolated bHLH domain is only marginally stable in solution, and partially folds upon dimerization. Binding to DNA promotes folding, stabilization, and protection from proteolysis of the bHLH domain. The Orange domain, on the contrary, is well folded in all conditions, forms stable dimers, and greatly increases protein resistance to thermal denaturation. The isolated proline-rich C-terminal region is mainly disordered in solution, and remains unstructured also in the full length protein. Measurements of binding constants show that HES-1 recognizes dsDNA synthetic oligonucleotides corresponding to several functional DNA targets with high affinity, but with relatively little specificity. We propose that order/disorder transitions in the different domains are associated not only with binding to DNA, but also with protein homo- and hetero-dimerization.

1. Introduction

The *HES1* gene, one of the seven human homologues of *Drosophila Hairy* and *Enhancer of Split* (*E(spl)*), encodes a DNA binding protein that acts as a transcriptional repressor [1–4]. As a general rule, HES-1 regulates tissue morphogenesis by maintaining precursor cells in an undifferentiated state [5]. The timely expression of HES-1 [6] is required for the correct development of many tissues such as the brain, eye and pancreas, and for the correct differentiation of hematopoietic cells. HES-1 was soon recognized as one of the main effectors of Notch signaling in mammals [7] and was recently shown to be required in quiescent cells to restart proliferation [8].

HES-1 belongs to the large family of transcription factors also known as bHLH (basic helix-loop-helix) proteins. These proteins can be further divided in four groups, homologues of *Drosophila* Hairy (including human HES-1 and -4), of *Drosophila* E(spl) (including human HES-2, -3, -5, -6, -7), the HEY group (including human HEY-1, HEY-2, and HEY-L) and the Stra13 group (including human DEC1 and DEC2) [4]. They all share the same domain architecture (Fig. 1)—a basic helix-loop-helix domain (bHLH) and an Orange domain (O)—and the same function, acting as transcriptional repressors, although with distinct specificities. Several DNA target sequences have been identified so far [1] including *HES1* own promoter [9], the promoter regions of the transcription factor Achaete Scute Homolog 1 (*ASH1*) [10], and of the cyclin-dependent kinase inhibitors $p21^{WAF-1/CIP1}$ and $p27^{kip1}$ [11,12]. While most bHLH proteins bind DNA as homo- or hetero-dimers at E-box sites (CANNTG), HES-1 (and HES-5) was reported to bind preferentially at N-box sites (CACNAG) [13] but sequence specificity has not been studied in detail. The basic region of HES-1 is characterized by a proline that is highly conserved in the Hairy and E(spl) groups, whereas a glycine is found at the same position in the related bHLH members of the HEY group.

The HLH domain has been proposed to promote protein dimerization. It may also contribute to the DNA binding interface, and to the correct orientation of the basic region in the major groove of the target DNA. HES-1 can form not only homo- but also hetero-dimers with other DNA binding proteins like Transcription Factor E2-α (E47/TCF3), inhibiting these transcription activators [13,14]. Formation of hetero-dimers by HES-1 with members of the same bHLH-O family like HES-6 [15] and HEY-2 [16] (also named HERP1/Hesr2/HRT2/CHF1/Gridlock) has also been reported. The functional effects have not been clearly established in all circumstances, but the hetero-dimer formation can be favored, in specific cases, with respect to the homo-dimer [17], and it has been suggested that the HES-1/HEY-2 hetero-dimer may be the functional repressor *in vivo*.

The Orange domain is required to inhibit transcription activation by Achaete-scute homolog 1 (ASH1) and E47 [11], but its precise function remains to be determined. It might contribute to the homo- and hetero-dimer interface, or be necessary for the recruitment of specific co-repressors, or both.

The C-terminal half of HES-1 has not been characterized yet, and contains a poly-proline region, a short conserved segment of yet unspecified role, also named as HC region, and a WRPW tetrapeptide

* Corresponding authors. Tel.: +39 040 3757354; fax: +39 040 226555.
E-mail addresses: pintar@icgeb.org (A. Pintar), pongor@icgeb.org (S. Pongor).

1570-9639/$ – see front matter © 2010 Published by Elsevier B.V.
doi:10.1016/j.bbapap.2010.08.010

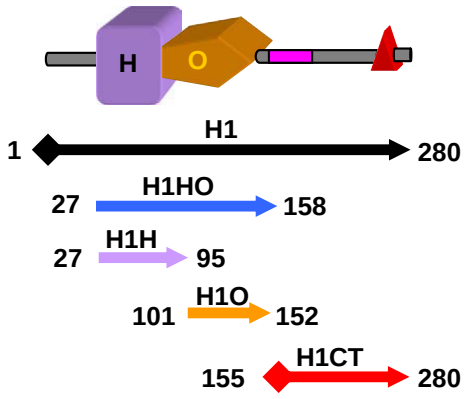


Fig. 1. Domain architecture of HES-1. Domain architecture of HES-1 and constructs studied in this work: H1, full-length HES-1; H1H, basic-helix-loop-helix domain; H1O, Orange domain; H1HO, bHLH-Orange tandem domains; H1CT, C-terminal region; the poly-proline segment is in magenta, the WRPW motif is shown as a red triangle, diamonds represent a His_6 tag. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

motif at the C-terminus, which is required to bind the transcriptional co-repressor Transducin-Like Enhancer protein 1 (TLE1/Groucho) [18]. It is not known if subtle variations in the C-terminal motif are used to recruit specific TLE proteins.

In the current study, we dissect HES-1 in its main structural and functional regions—the bHLH domain, the Orange domain, the C-terminal half—and characterize in detail the biochemical and spectroscopic properties of these modules either as isolated domains or in combination as in the native protein (Fig. 1). Furthermore, we report the K_d values for the complexes between the different HES-1 constructs and a series of nine dsDNA oligonucleotides with sequences based on known targets *in vivo*.

2. Materials and methods

2.1. Construction of expression plasmids

The different constructs of HES-1 used in this study cover the basic helix-loop-helix domain (H1H), the Orange domain (H1O, prepared by solid phase peptide synthesis, see below), the basic helix-loop-helix-Orange tandem domains (H1HO), the C-terminal half containing the proline rich region and the WRPW motif (H1CT), and the full-length protein (H1) (Fig. 1).

The H1 DNA was amplified by PCR from the IMAGE cDNA clone 4749611 (ImaGenes GmbH, Berlin, Germany) which encodes a 277 residue protein (Swiss-Prot: Q8IXV0_HUMAN) differing from the 280 residue HES1_HUMAN only for three residues in the N-terminal region (^{16}ATP^{18}) (Fig. S1 in the Supporting Material). A synthetic gene encoding the H1HO domain (residues 27–158) was assembled and amplified in a two-step PCR from a series of 18 oligonucleotides optimized for expression in *E. coli* (Sigma Genosys). The DNA encoding the H1H domain (residues 27–95) was amplified by PCR from the synthetic H1HO gene construct. The DNA encoding the H1CT region of HES-1 (residues 154–280) was amplified by PCR from the full-length H1 construct. In all of the constructs, the forward and the reverse primers carry an *Nde*I and a *Bam*HI restriction site, respectively. In addition, the forward primer used to obtain the H1 and H1CT DNA also contains a polyhistidine coding sequence (Fig. S1 in the Supporting Material). All of the PCR products were digested with *Nde*I and *Bam*HI, and directionally subcloned into a pET-11a expression vector (Novagen). *E. coli* DH5α cells were transformed, and selected on LB plates containing 100 μg/ml ampicillin. Correct sequences of selected clones were verified by automatic DNA sequencing (BMR Genomics, Padua, Italy).

2.2. Protein expression and purification

Recombinant proteins were expressed in *E. coli* BL21(DE3) cells containing the pLysS plasmid (Novagen). Bacteria were grown at 37 °C in LB medium containing 100 mg/l ampicillin and 25 mg/l chloramphenicol to a density of ~1 OD_{600} units and protein expression induced with 1 mM IPTG for 3 h. Cells were harvested by centrifugation and pellets frozen at −70 °C. The H1 and H1CT proteins, carrying a His_6 tag, were purified under native conditions by nickel-ion affinity chromatography (see Supporting Material for details). The H1HO and H1H proteins were purified in two steps, by ion exchange chromatography followed by reverse-phase high-pressure liquid chromatography (RP--HPLC) (see Supporting Material for details). All the purified proteins were analyzed by SDS-PAGE and liquid chromatography-mass spectrometry (LC-MS) to assess their purity and correct molecular weight.

2.3. Peptide synthesis

The peptide (52 amino acid long) corresponding to H1O (residues 101–152 of human HES-1, with C117, C128, and C146 mutated to alanines) was automatically synthesized on solid phase (TentaGel™ S PHB-Ala-Fmoc resin, 0.05 mmol scale) using Fmoc/tBu chemistry and purified by RP-HPLC (see Supporting Material for details). The purified peptide fractions were analyzed by LC-MS to verify purity and molecular mass, pooled, and freeze-dried.

2.4. Size exclusion chromatography

The protein samples were dissolved in the elution buffer (50 mM NaP pH 7.4, 150 mM NaCl, 10% glycerol, 2 mM EDTA), loaded onto a Superdex 75 column (Pharmacia) and eluted in the same buffer. The apparent molecular mass of the different constructs was estimated from a calibration carried out with the following molecular standards: bovine serum albumin (67 kDa), ovalbumin (43 kDa), myoglobin (17 kDa), ribonuclease A (13.7 kDa) and aprotinin (6.5 kDa).

2.5. Circular dichroism (CD)

Samples for CD spectroscopy were prepared dissolving the freeze-dried proteins in buffer (20 mM sodium phosphate, pH 7.4). Protein concentrations for the H1, H1O, H1HO, and H1C samples were determined by UV absorbance at 280 nm using the calculated ε value of 15,470, 1490, 2980, and 13,980 M^{-1} cm^{-1}, respectively. As the H1H construct contains no aromatic residues, protein concentration was estimated by amino acid analysis (Alta Bioscience, Birmingham, UK). CD spectra were recorded on a Jasco J-810 spectropolarimeter using jacketed quartz cuvettes of 0.1–10 mm pathlength. Typically, 1–5 scans were acquired for each spectrum in the range 190–250 nm at a scan rate of 20 nm/min. Thermal denaturation studies were carried out using a water bath connected to the jacketed quartz cuvettes. Curve fitting was achieved using the Levenberg–Marquardt least squares minimization algorithm as implemented in Kaleidagraph™ (Synergy Software). The sigmoid function, $y = M2 + (M1 - M2)/(1 + \exp((M3 - T)/M4))$ where M2 is the initial value, M1 the final value, M3 the inflection point, M4 the width of the transition, and T the temperature, was fitted to thermal denaturation (θ_{222}) data of H1H. For thermal denaturation of H1HO and H1O, the generalized logistic curve, $y = M2 + (M1 - M2)/(1 + M5 * \exp((M3 - T)/M4))^{1/M5}$ where M2 is the initial value, M1 the final value, M3 the inflection point, M4 the width of the transition, M5 the asymmetry of the curve, and T the temperature, was fitted to θ_{222} data. To estimate the dissociation constant K_d of the dimer, the quadratic function $y = M1 + (M2 - M1) * (\mathrm{sqrt}(M3 * (M3 + 8 * c)) - M3)/(4 * c)$ where M1 and M2 are the ellipticity at 222 nm of the monomer and the dimer, respectively, M3 the K_d and c the total protein concentration, was fitted to θ_{222} data. Mean residue ellipticity (MRE, deg cm^2 $dmol^{-1}$ $residue^{-1}$) was calculated

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010

from the baseline-corrected spectrum. A quantitative estimation of secondary structure content was carried out using CDSSTR run from the DichroWeb server [19]. Helical content was also estimated from the ratio of ellipticity at 222 and 208 nm ($\theta_{222}/\theta_{208}$).

2.6. Electrophoretic mobility shift assays (EMSA)

The oligonucleotides used in gel shift experiments were purchased from Sigma Genosys (HPLC grade). Each single-stranded oligonucleotide was 5′-end labeled with [γ-^{33}P]ATP (Perkin Elmer) using T4 polynucleotide kinase (New England Biolabs), purified, and annealed with a 100-fold excess of the cold complementary strand. The DNA duplex obtained was equilibrated in binding buffer at a final concentration of 40 pM. After addition of the purified protein, binding mixtures were incubated for 30 min at room temperature and subjected to non-denaturing polyacrylamide electrophoresis (see Supporting Material for details). Gels were fixed, dried and evaluated by the Cyclone™ Phosphor Storage System (Packard). The fraction of bound DNA, f, defined as $[DNA_{bound}]/([DNA_{bound}]+[DNA_{free}])$ was plotted vs. the total concentration, c, of the protein dimer, and K_d estimated from curve fitting of the function $f=c/(c+K_d)$ using KaleidaGraph™ (Synergy Software).

2.7. Limited proteolysis

Aliquots of purified proteins (20 μg H1H or 40 μg H1HO) were dissolved in 50 μl 2× digestion buffer (40 mM Tris–HCl, 0.3 M NaCl, 10 mM DTT, 10 mM $CaCl_2$) and each sample was then divided in two aliquots. The annealed DNA oligonucleotide (12mer or 20mer, dissolved in 25 μl water) was added in slight molar excess (1.33×) to one of the aliquots while the same volume of water was added to the other aliquot. The two mixtures were then incubated at RT for 1 h. After incubation, trypsin (0.1 μg, 1:100 w/w with respect to the protein) (sequencing grade, Promega, Madison, WI) was added and the samples incubated at 37 °C. After 1, 2, 5, 10, 30, and 60 min aliquots of 5 μl were quenched with 1 μl of 10% acetic acid. The quenched reaction mixture (1 μl) was then mixed with 9 μl of MALDI matrix (10 mg/ml α-Cyano-4-hydroxycinnamic acid in 70% acetonitrile/0.1% trifluoroacetic acid), spotted (1 μl) on the MALDI plate and analyzed in reflector positive ion mode on a 4800 MALDI-TOF/TOF Proteomics Analyzer (Applied Biosystems).

3. Results

3.1. The bHLH domain is marginally stable

At odd with secondary structure predictions, H1H displays no or little secondary structure in water (Fig. S2 in the Supporting Material), as evidenced by the negative CD band at 200 nm and by the very low ellipticity at 190 nm, which are typical features of disordered polypeptides. The increase of ionic strength through the addition of KF progressively induces the formation of α-helical secondary structure, as proved by the shift of the negative band to 208 nm, the formation of an additional negative band at 222 nm, and the formation of a strong positive band at 190 nm. However, when H1H is initially dissolved in a relatively low concentration (20 mM) phosphate buffer, a “native” CD spectrum is obtained and a further increase in the ionic strength through the addition of KF up to 150 mM does not induce any additional change in the CD spectrum (Fig. S2B in the Supporting Material). The relatively low concentration of phosphate buffer (20 mM) sufficient to promote folding of H1H suggests that different ions can have different effects on the structure of H1H. Indeed, the large phosphate ion, with its capability of forming multiple hydrogen bonds, might stabilize the helical structure of H1H, possibly mimicking the phosphate moieties of the DNA backbone. Folding of H1H is actually promoted also by binding to DNA (see Section 3.6).

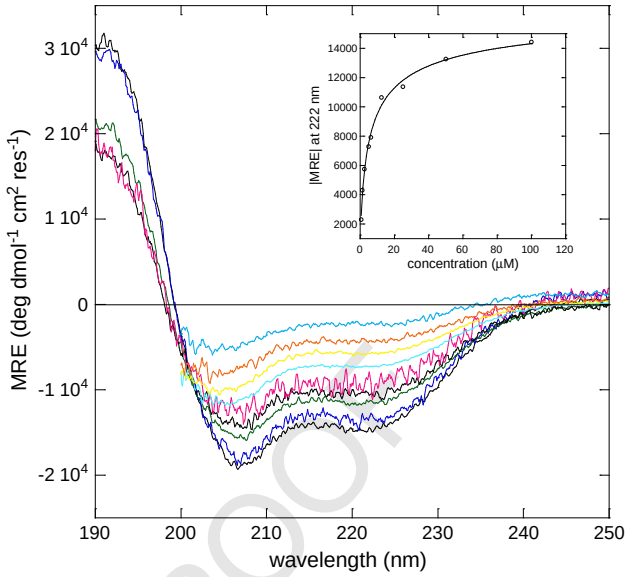


Fig. 2. Effect of concentration on H1H. Far-UV CD of H1H at different concentrations (100, 50, 25, 12.5, 6.3, 5.0, 2.5, 1.2 μM); inset: MRE at 222 nm plotted vs. protein concentration (see Materials and methods for details on curve fitting).

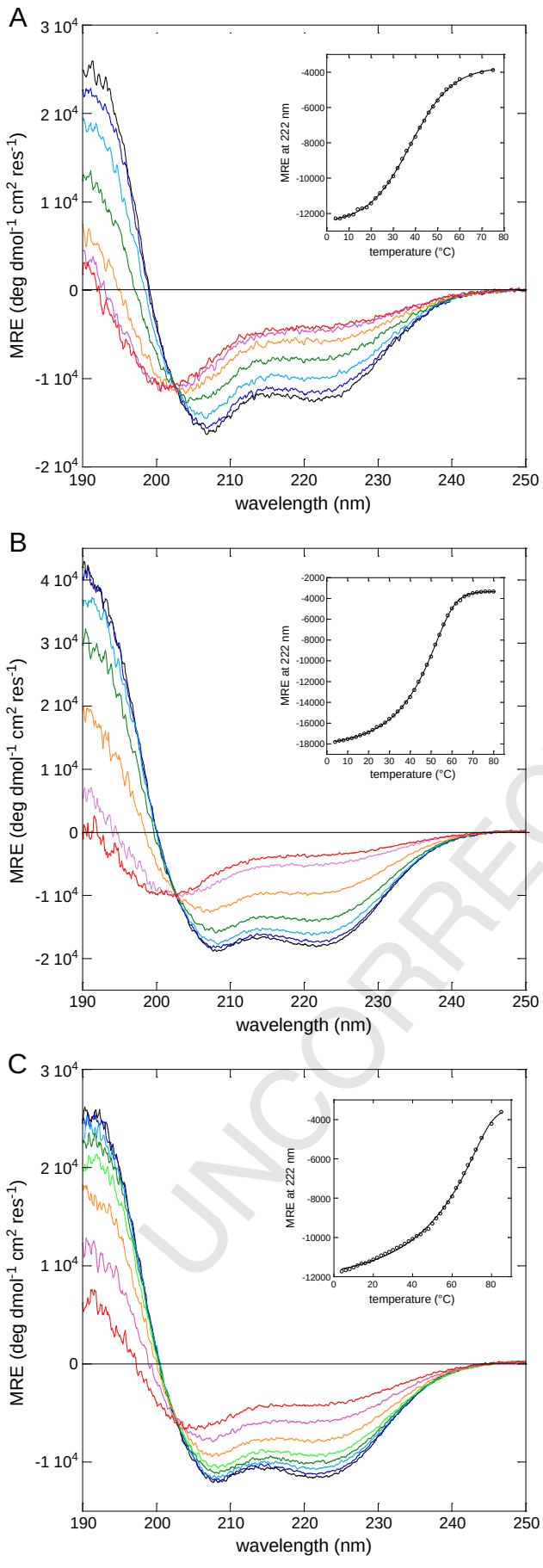
From these results, we hypothesized that folding of H1H is associated with the dimerization of the bHLH domain. As dimerization is a concentration-dependent process, we recorded far-UV CD spectra at constant ionic strength (in buffer) and at different protein concentrations (Fig. 2), and plotted the ellipticity at 222 nm as a measure of the helical content. The normalized intensity of the CD bands displays a strong dependence on protein concentration. Furthermore, also the shape of the far-UV CD spectrum changes, with a progressive shift of the negative band from 208 to 203 nm at lower protein concentrations, which suggests a shift towards a disordered conformation occurring even in buffer. From the plot of ellipticity at 222 nm it was possible to estimate the K_d of the dimer (8.4 ± 2.1 μM). The dimeric state of H1H in high ionic strength conditions was confirmed by size exclusion chromatography (Fig. S3 in the Supporting Material). Although the apparent MW estimated from size exclusion chromatography (23.5 ± 1.9 kDa) is higher than the value expected for a dimer (16.4 kDa), this result is consistent with an only partially globular state of the H1H dimer, where the N-terminal region corresponding to the predicted helices b1 and b2 (residues 27–47) is mainly disordered in solution in the absence of DNA. The disorder in the N-terminal region is supported by the secondary structure content obtained by deconvolution of the far-UV CD spectra (Table 1 and Fig. S4 in the Supporting Material) and by the sensitivity of this region to limited proteolysis (see below), and would lead to an increase in the hydrodynamic radius of the protein, hence to an increase in the apparent MW estimated by size exclusion chromatography. The stability of H1H, as measured by

Table 1 t1.1
Secondary structure. Predicted helical content (α pred., %) and secondary structure content (α, helix; β, β-strand; T, turn; c, coil, %) calculated by deconvolution of the far-UV CD spectra. For predictions, a lower and upper bound is given; the lower bound considers only residues predicted to be helical with a high score, the upper bound all residues predicted to be helical; percentages are calculated assuming that the basic region is not folded; percentages reported in parenthesis assume that also the basic region is folded.

t1.2

	α pred.	α	β	T	c	t1.3
H1	20–34 (27–40)	25	17	16	42	t1.4
H1HO	43–72 (56–85)	45	12	16	28	t1.5
H1H	38–56 (62–81)	54	9	11	26	t1.6
H1O	60–71	61	8	8	22	t1.7
H1CT	0	5	22	16	57	t1.8

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010



thermal denaturation followed by far-UV CD, was found to be marginal. Increasing the temperature from 4 to 75 °C, a clear change in the CD spectrum can be detected already at 40 °C, with a decrease in the intensity of the bands at 190 and 222 nm, and a shift towards lower wavelengths of the band at 208 nm (Fig. 3A). Over 50 °C, the CD spectrum is that of a mainly unfolded protein. The transition was found to be reversible, as judged from the far-UV CD spectra recorded after cooling, and by the recovery of the ellipticity at 222 nm (>95%). Plotting ellipticity at 222 nm vs. temperature, we obtained a T_m (inflection point) of 37.7 ± 0.1 °C, with a half-width of the transition of 9.4 ± 0.1 °C. H1H is therefore metastable, with a denaturation temperature that is, in the experimental conditions used, very close to the physiological temperature. We can conclude that H1H is in equilibrium between a disordered monomer and a partially folded dimer with a low thermal stability. Accordingly, the CD spectra of H1H are strongly affected by three factors: ionic strength of the solution, and, most importantly, temperature and protein concentration. Higher salt concentration, lower temperature, and higher protein concentration shift the equilibrium towards the partially folded dimer, where the N-terminal and basic regions remain mainly disordered, and the HLH motif acquires helicity.

3.2. The Orange domain is a stable helical dimer

The H1O protein, which encompasses the Orange domain, displays a completely different behavior. Even in low ionic strength conditions H1O is folded, with a secondary structure content consistent with a mainly helical Orange domain, and the far-UV CD spectrum is not affected by an increase in ionic strength obtained by addition of either KF or KCl (Fig. S5 in the Supporting Material). In buffer, we did observe a dependence of the far-UV CD spectrum intensity upon protein concentration (data not shown), but the change is less remarkable than that of H1H, the change in the shape of the spectrum, with a shift of the negative band from 208 to 206 nm is hardly detectable, and from the plot of ellipticity at 222 nm vs. protein concentration it was not possible to calculate a K_d. These results are consistent with an equilibrium, similar to that postulated for H1H, between an unfolded monomer and a folded dimer where the dimer is relatively stable and the K_d is out of the range measurable by CD. H1O was found to be indeed more stable than H1H. Thermal denaturation of H1O was followed by CD, recording far-UV CD spectra (Fig. 3B) and plotting ellipticity values at 222 nm between 4 and 80 °C. Also in this case the transition is reversible, and from ellipticity values at 222 nm an inflection point (T_m) at 52.5 ± 0.1 °C was calculated. The apparent MW of H1O determined by size exclusion chromatography is 13.0 kDa, which confirms the dimeric state of H1O (calculated MW for the dimer: 11.2 kDa) (Fig. S4 in the Supporting Material).

3.3. The C-terminal half is mainly disordered

Apart from the bHLH and the Orange domains, HES-1 presents a rather large C-terminal half that is made of a proline-rich region (residues 156–202) and a C-terminus that is carrying the WRPW motif. This segment is predicted to be mainly in a random coil conformation (Fig. S1 in the Supporting Material) and the recombinant protein, H1CT, corresponding to this region actually shows the characteristic CD spectrum of a disordered protein, with a negative band close to 200 nm, the absence of any negative band at 222 nm, and a negative ellipticity at 190 nm (Fig. S4 in the Supporting Material). These results are further

Fig. 3. Thermal stability. Far-UV CD of (A), H1H (50 μM) at different temperatures (10 °C, black; 20 °C, blue; 30 °C, light blue; 40 °C, green; 50 °C, orange; 60 °C, magenta; 70 °C, red), (B), H1O (97 μM) (10 °C, black; 20 °C, blue; 30 °C, light blue; 40 °C, green; 50 °C, orange; 60 °C, magenta; 70 °C, red), (C), H1HO (50 μM) (10 °C, black; 20 °C, blue; 30 °C, light blue; 40 °C, dark green; 50 °C, light green; 60 °C, orange; 70 °C, magenta; 80 °C, red); insets: MRE at 222 nm vs. temperature; data were recorded at steps of 2 °C (see Materials and methods for details on curve fitting). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010

confirmed by NMR spectra, which show no significant dispersion in the chemical shifts of the backbone NHs (data not shown). From size exclusion chromatography, an apparent MW of 15.8 kDa can be estimated (calculated: 13.6 kDa), which is consistent with the expanded hydrodynamic radius expected for a natively disordered protein (Fig. S3 in the Supporting Material). Although it can be speculated that the C-terminal region might fold in the context of the full-length native protein, secondary structure content calculated by deconvolution of the CD spectra of the different constructs suggest that this is not the case (Table 1). This is further confirmed by thermal denaturation followed by far-UV CD, which shows that the C-terminal half is not contributing to the overall stability of HES-1 (Fig. S6 in the Supporting Material).

3.4. Secondary structure and oligomeric state of HES-1

The secondary structure content of the full-length HES-1 (H1), of the bHLH-Orange tandem domains (H1HO), of the bHLH domain (H1H), of the Orange domain (H1O) and of the C-terminal region (H1CT) was calculated by deconvolution of the far-UV CD spectra (Table 1 and Fig. S4 in the Supporting Material). CD data confirm the helical conformation of the bHLH and Orange domains, whereas the C-terminal region is mainly disordered. The helical content predicted for H1, H1H, and H1HO is systematically larger than that found by CD spectroscopy. However, a very good agreement between predictions and experimentally determined values is found if the basic region, despite its propensity to adopt a helical structure, is assumed to be in a random coil conformation in the absence of DNA.

Results from size exclusion chromatography of H1 and H1HO suggest a dimeric state of HES-1 (Fig. S3 in the Supporting Material). The apparent MW estimated for H1 and H1HO are 76.6 ± 8.1 and 34.7 ± 4.2 kDa, respectively. These values are larger than those calculated for the respective dimers (60.8 and 30.2 kDa) but are consistent with the only partially globular nature of these proteins. CD data show that the basic region is mainly disordered in the absence of DNA and that in H1 the entire C-terminal half of the protein is in a random coil conformation. The presence of these disordered regions leads to an increase in the effective hydrodynamic radius measured by size exclusion chromatography.

The thermal stability of the structured portion of HES-1, which includes the bHLH and the Orange domain (H1HO) is remarkable. From far-UV CD spectra (Fig. 3C), a significant decrease in the intensity of the bands at 190, 208, and 222 nm, together with a shift towards shorter wavelength of the band at 208 nm can be detected only at temperature of 60 °C or higher. The plot of ellipticity at 222 nm recorded between 4 and 85 °C does not show a clear cut transition but rather two phases, the first between 4 and 45 °C, the second between 45 and 85 °C, where ellipticity is progressively decreasing. Curve fit of thermal denaturation data leads to an inflection point (T_m) at 71.5 ± 0.8 °C. Cooling of the sample allows for a nearly complete (~90%) recovery of ellipticity at 222 nm and of the far-UV CD spectrum, showing that the unfolding of H1HO is largely reversible, despite the presence of three cysteines in the Orange domain. The full-length protein (H1) behaves in a different way with respect to thermal denaturation (Fig. S6 in the Supporting Material). Thermal denaturation of H1 is not fully reversible, as only ~50% of the ellipticity at 222 nm is recovered after a heating/cooling cycle. Although the CD spectra (Fig. S6 in the Supporting Material) suggest a remarkable thermal stability of H1, the decrease in the ellipticity at 222 nm with temperature is not accompanied by a significant change in the shape of the CD spectrum. Accordingly, the plot of ellipticity at 222 nm is not sigmoidal and the T_m could not be calculated. This suggests that the C-terminal disordered region, while not contributing to the overall stability of the protein, may induce aggregation or precipitation of H1 upon heating and prevent correct refolding of the partially unfolded forms.

3.5. HES-1 binds DNA with high affinity but low specificity

Affinity of the bHLH domain (H1H), of the bHLH-Orange tandem domains (H1HO), and of the full-length protein (H1) towards a series of dsDNA synthetic oligonucleotides of equal length (20mer) representing biologically relevant targets were evaluated by electrophoretic mobility shift assay (EMSA) and are reported in Table 2. These targets were selected among the DNA sequences recognized by HES-1 *in vivo* and cover a variety of consensus binding sites. N1, N2, and N4 represent three N-box sites (CACNAG) in the promoter of the human *HES1* gene. They differ for the base N (A, G, C, respectively) in the consensus recognition site and for the flanking regions. N2 can also be classified as a type C site (CACGNG). EA is a mutant of N1 where the N-box was changed into an E-box (CANNTG), subtype A (CANCTG). The hp27 and hp21 sequences are found in the promoter of the cyclin-dependent kinase inhibitors $p27^{kip1}$ and $p21^{WAF-1/CIP1}$, respectively. They share an identical binding site (type C, CACGNG) but different flanking regions. The hASH1 oligonucleotide corresponds to the HES-1 binding site identified in the promoter region of the transcription factor Achaete Scute Homolog 1 (*ASH1*). The sequence does not correspond to a canonical N- or E-box and can be defined as an atypical class C site. A second binding site in the same promoter is represented by hASH1c (class C). The hASH1c* oligonucleotide is a mutant of hASH1c where the class C site was changed into a palindromic E-box, class B. With few exceptions, most of the K_d values are in the 2–10 nM range (Table 2 and Fig. S7 in the Supporting Material). Although it has been reported that HES-1 binds preferentially to N-box rather than to E-box consensus sites, we are not able to confirm this conclusion. Especially considering the K_d values obtained for H1H and H1HO, which gave more reproducible results, there is not a clear preference for a specific target sequence. Even for mutated sequences, such as in the EA and hASH1c* oligonucleotides, which were originally designed as negative controls, the measured K_d are in the nM range and, taking into account the experimental error associated with these experiments, not significantly different from those measured for the canonical target sequences. In conclusion, HES-1, although binding all the examined sequences with relatively high affinity, did not display the expected specificity. This was further confirmed by binding experiments carried out in the presence of a poly(dI–dC) as a nonspecific competitor (Fig. S8 in the Supporting Material).

Table 2
DNA binding. Synthetic dsDNA oligonucleotide (20mer) identifier, classification of the binding site (N, N-box: CACNAG; E, E-box: CANNTG; A, class A: CANCTG; B, class B: CANGTG; C, class C: CACGNG; class A and B sites are subtypes of the E-box; the N-box and class C sites are partially overlapping), sequence of the dsDNA oligonucleotide (binding site in bold underlined), K_d values (nM) and standard deviations determined by Electrophoretic Mobility Shift Assay (EMSA).

	Class	Sequence	H1H	H1HO	H1
N1	N	5′-GCCAGACCTTGTGCCTGGCG CGGTCTG7GAACACGGACCGC-5′	3.9 ± 1.3	1.6 ± 0.7	2.8 ± 0.6
EA	E/A	5′-GCCAGACCAGGTGCCTGGCG CGGTCTGGTCCACGGACCGC-5′	4.0 ± 1.6	2.0 ± 0.9	12.6 ± 0.9
N2	N/C	5′-CGCGGTCCACGAGCGGTGCC GCGCCAGGTGCTCGCCACGG-5′	4.2 ± 1.7	2.8 ± 1.7	6.3 ± 1.8
N4	N	5′-TCCGGAGCTGGTGCTGATAA AGGCCTCGACCACGACTATT-5′	6.0 ± 1.9	3.5 ± 1.4	8.4 ± 0.9
hp27	C	5′-CAGCAGTCACGCGACCAGCC GTCGTCAGTGCGCTGGTCGG-5′	2.0 ± 0.9	2.2 ± 1.3	4.5 ± 1.3
hp21	C	5′-CCTGCAGCACGCGAGGTTCC GGACGTCGTGCGCTCCAAGG-5′	3.8 ± 1.2	2.7 ± 1.5	6.8 ± 0.2
hASH1	-	5′-CCAGGCGCACGCACTGCAAC GGTCCGCGTGCGTGACGTTG-5′	3.7 ± 0.4	1.8 ± 1.1	3.0 ± 0.7
hASH1c	C	5′-AGTCCGGCACGCGCCAGGCG TCAGGCCGTGCGCGGTCCGC-5′	2.3 ± 0.9	1.2 ± 0.4	3.0 ± 0.8
hASH1c*	E/B	5′-AGTCCGGCACGTGCCAGGCG TCAGGCCGTGCACGGTCCGC-5′	1.8	2.6	5.9

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010

3.6. Binding to DNA promotes folding

The coil-helix transitions occurring upon binding to dsDNA were monitored by CD. At low protein concentration, binding of H1H (Fig. 4A) to the synthetic 20mer dsDNA oligonucleotide N1 is associated with a significant change in the CD spectrum, with a shift of the negative band from 205 to 208 nm and a strong variation in the intensity of the bands at 190 and 222 nm. This result is in support of a shift of the equilibrium from a largely disordered state to a helical conformation of H1H, promoted by binding to DNA. At higher protein concentrations, H1H is mainly present as a folded dimer, and the variations in the CD spectrum upon binding to DNA are less evident, but still detectable (Fig. S10 in the Supporting Material). Binding to DNA is also accompanied by a drastic increase in the thermal stability of H1H. Far-UV CD spectra of H1H bound to DNA did not show any significant change between 10 and 60 °C (Fig. S9 in the Supporting Material), at odd with the only marginal thermal stability of the free H1H (Fig. 2A). Binding of H1HO (Fig. 4B) to the same dsDNA oligonucleotide is associated with a more subtle variation in the CD spectrum, namely the change in the relative intensity of the bands at 208 and 222 nm, which suggests a slightly higher helical content in H1HO when bound to DNA. This is consistent with the HLH domain being mainly folded in the H1HO protein even in the absence of DNA. The folding of the basic region upon binding to DNA was confirmed by limited proteolysis experiments followed by analysis of the fragments by MS (Fig. 5). In the absence of DNA, H1H is rapidly cleaved by trypsin at several positions in the regions b1 and b2, with the largest fragment corresponding to helices h1 and h2. In the presence of DNA, H1H is overall well protected from cleavage and larger fragments, including the intact protein, can be identified. Similar results were obtained for H1HO (data not shown). Although limited proteolysis experiments do not provide a direct evidence of a conformational change, in combination with CD they are in support of a coil-helix transition of the b2 region occurring upon binding to DNA. Furthermore, the partial resilience of b1 to proteolytic cleavage suggests that also this region may be involved in DNA binding.

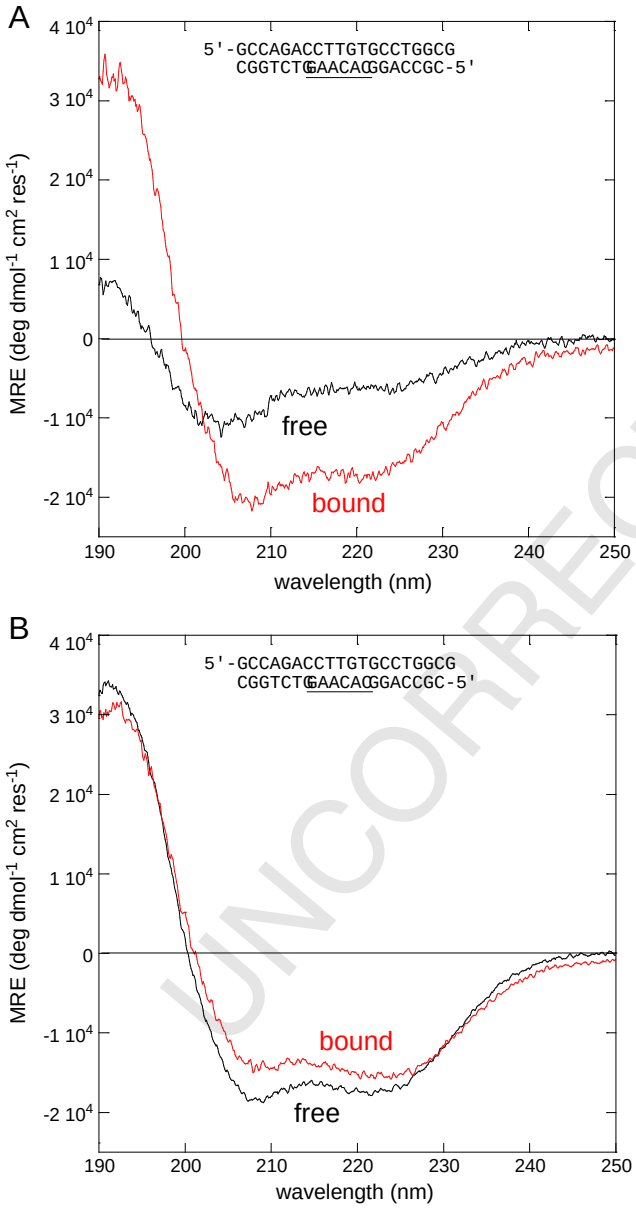


Fig. 4. Folding upon binding to DNA. Far-UV CD spectra of (A), H1 (5 μM) and (B), H1HO (4 μM) in the presence of 1 equivalent of N1 synthetic dsDNA oligonucleotide. The CD spectrum of the free protein is shown in black, the CD spectrum of the protein bound to DNA, obtained by subtracting the spectrum of the dsDNA from the spectrum of the complex, in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

The overall domain organization of human HES-1 emerging from this study shows a dimeric protein with a mainly disordered N-terminal region (including b1 and b2) that is undergoing a coil-helix transition upon binding to DNA, a globular, helical region formed by the HLH and the Orange domains, and an intrinsically disordered C-terminal half. Our results on H1H confirm early studies carried out on a synthetic peptide corresponding to the bHLH domain of Deadpan, the *Drosophila* homologue of human HES-1 [20]. It was shown by CD spectroscopy that this peptide, which has 72% sequence identity with residues 33–96 of human HES-1, is mainly unstructured in low ionic strength conditions, acquires helical structure upon addition of KCl and significantly increases its helical content upon binding to DNA. It was also found by analytical ultracentrifugation that this peptide is a dimer in high salt conditions.

It has been proposed that the HLH domain is required for dimerization of the transcription factors of this large family. We found, on the other hand, that the bHLH domain of HES-1 is only marginally stable at physiological temperature and in the μM concentration range. In these conditions, H1H is likely to be in equilibrium between a monomeric unfolded form and a dimeric folded form. On the contrary, the Orange domain is more stable and seems to be better suited to provide the driving force for dimerization of the full-length protein. The only structure currently available for an Orange domain—the crystal structure of the Orange domain from human HEY-1 (PDB: 2DB7)—reveals the presence of a homodimer where the two monomers are covalently linked by a disulfide bond formed by the single endogenous cysteine present in each monomer. Because the basic conditions (pH 9.2) in which the protein was crystallized favor the rapid oxidation of cysteines, it could be questioned whether the dimer formation is authentic. Our results suggest that homodimerization could be indeed a general feature of the Orange domain. To avoid possible artifacts arising from the three cysteines present in the Orange domain of HES-1, the H1O polypeptide used in this study had the three cysteines mutated to alanines (C117A/C128A/C146A). The rationale for replacing all three cysteines with alanines rather than, for example, with serines was based on multiple sequence alignments (data not shown) and on the assumption that the cysteines are buried either at the interface between the two predicted helices or at the interface between the two units of the dimer, as in the crystal structure of the HEY-1 Orange domain.

The biological role of the transcriptional repressors of the HES family is strictly related to their networking potential, i.e. their capability of forming heterodimers with other proteins. These can be members of the

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010

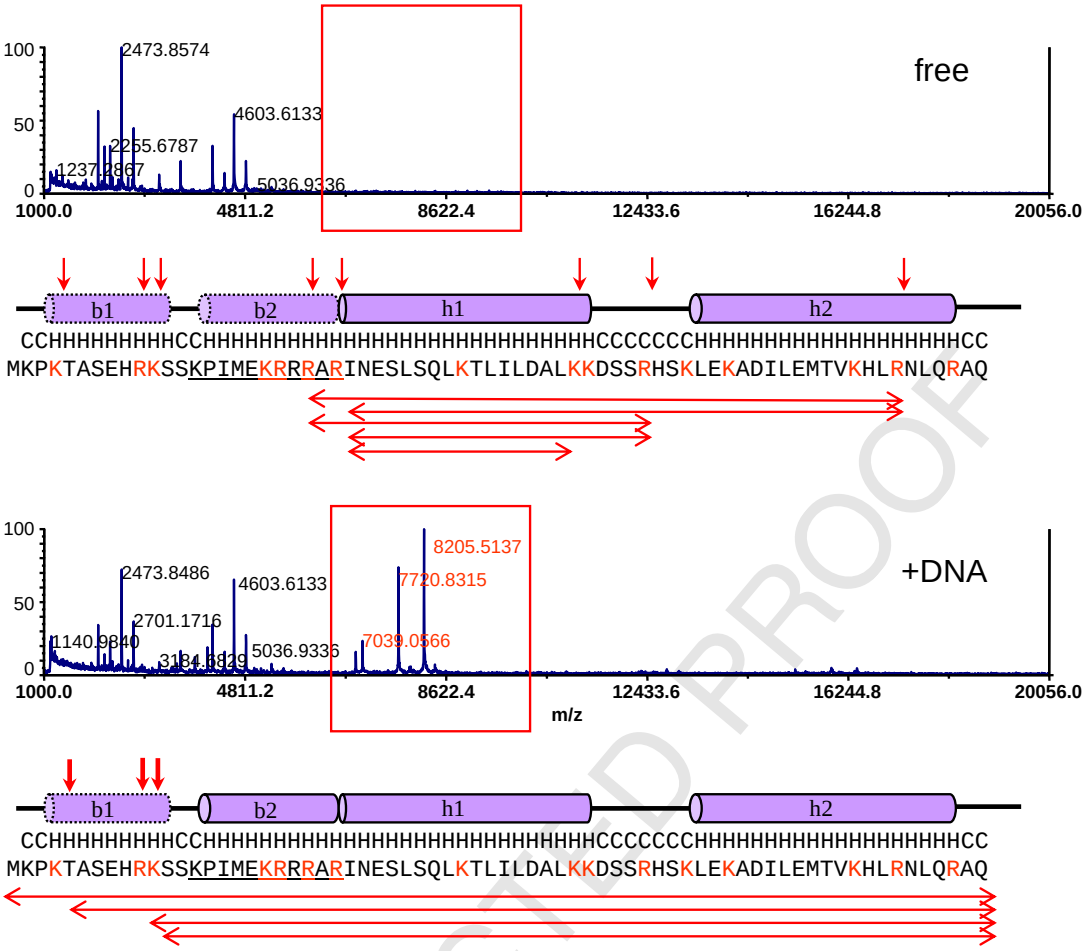


Fig. 5. Limited proteolysis. Limited proteolysis of H1 alone (top) and in the presence (bottom) of the synthetic N1 dsDNA oligonucleotide (20mer); the amino acid sequence and the predicted secondary structure are also shown; trypsin potential cleavage sites are in red in the amino acid sequence; determined cleavage sites are indicated by red arrows and the corresponding fragments identified by MS as double arrows under the sequence; the red frame highlights the fragments that are protected in the presence of DNA but not in the free protein. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

same family such as HES-6, members of the related HEY family such as HEY-1 and -2, other transcription factors, or inhibitors of DNA binding such as ID proteins. The combinatorial optimization of the network would lead to the maximum number of potential targets using a limited number of proteins. In principle, heterodimer formation can be driven by mass effect through variations in the expression levels occurring

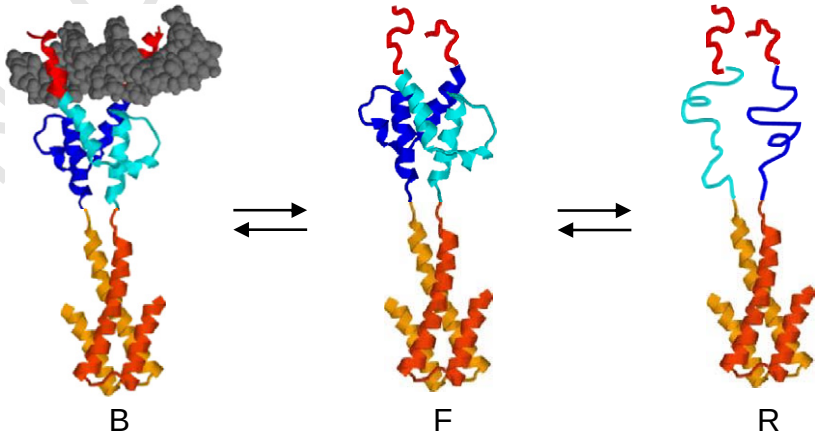


Fig. 6. Structural states of HES-1. A cartoon model of HES-1 in its DNA bound (B), free (F) and relaxed (R) states. The basic region is in red, the HLH domain in cyan (chain 1) or blue (chain 2), the Orange domain in orange (chain 1) or yellow (chain 2). The two chains can be identical (homodimer) or different (heterodimer). In the B state the protein, including the basic region, is fully folded; in the F state the basic region is mainly disordered but the HLH and Orange domains are folded, thus locking the dimeric form; the F state is in equilibrium with the R state, where the HLH domain is at least partially unfolded; in the R state the temporary increase in the K_d of the dimer facilitates exchange with other bHLH proteins. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010

during time, by differences in the K_d values of the homo- and heterodimer (different specificities) modulated by subtle differences in the amino acid sequence of the interface, or by post-translational modifications that can modulate the K_d. In any case, to allow for exchange to occur, the K_d values should be of the same order of magnitude of the physiological concentration of these transcription factors, which is probably in the μM range. We speculate that the HLH domain, with its marginal stability especially at physiological temperatures, might work as a lock and release mechanism, where the equilibrium between the folded and the unfolded form would temporary lock and release, respectively, the dimer, temporary changing its K_d. This mechanism would allow for exchange of the monomer units to yield heterodimers with different specificities (Fig. 6). This hypothesis will have to be confirmed by more detailed studies on the structural and dynamic properties of H1H and H1HO in solution. Folding of the basic region associated with binding to specific DNA targets would fix the structure in its most stable state, as shown by the drastically increased thermal stability of H1H when bound to DNA.

Determination of the binding affinities towards different synthetic dsDNA oligonucleotides covering different classes of binding sites and corresponding to *in vivo* targets did not confirm the expected specificity. It cannot be ruled out that this may be due to technical limitations. First, EMSA is not an equilibrium technique in rigorous terms, as it is assumed that no significant dissociation of the DNA/protein complex is occurring during the electrophoresis run. Second, curve fitting of the experimental data for the estimation of K_d values assumes that a binding model is selected. For homogeneity, we used the dimer mode of binding, but this assumption might not be verified in all conditions. However, the high affinity but relatively low specificity of binding of basic helix-loop-helix proteins such as MASH-1 [21] and MyoD [22] towards DNA targets had already been remarked, in contrast with the specificity of their biological effects, and HES-1 seems to behave in the same way. A possible explanation is that specificity of HES-1 *in vivo* is achieved through the formation of heterodimers with other bHLH proteins, either of the same group, or of the related HEY family. Indeed, it can be expected that recognition of symmetric (palindromic) DNA sequences can be effectively attained by symmetric homodimers, whereas asymmetric DNA sequences are more likely to be recognized by asymmetric heterodimers. This simple explanation is however questionable, as it was proposed that the bHLH homodimer of *Drosophila* deadpan can recognize asymmetric dsDNA sequences through unique interactions [23]. Furthermore, specificity may also arise from contacts with the DNA regions flanking the consensus binding site. *In vivo*, the molecular mechanisms of transcriptional repression or activation are likely to be much more complex, and not limited to the specific protein–DNA recognition. Both DNA binding-dependent and DNA binding-independent mechanisms of transcriptional control are possible, including the recruitment of different co-factors that would lead to the formation of DNA-bound multiprotein complexes. The recent identification of a direct interaction between HES-1 and the Fanconi anemia core complex [24], a macromolecular assembly of more than a dozen proteins involved in DNA repair, supports this view. Furthermore, post-translational modifications such as phosphorylation have been demonstrated to play a role in transcriptional activity of HES-1 [25,26], and may modulate both protein–DNA and protein–protein interactions. Other post-translational modifications, such as lysine acetylation and Ser/Thr *O*-GlcNAcylation, although not yet reported in HES-1, have been identified in other bHLH proteins, and are likely to contribute to transcriptional regulation. Finally, time- and context-dependent 1expression of transcriptional activators and inhibitors can also contribute to the specificity of the biological response.

5. Conclusions

We provide here the first biochemical characterization of human HES-1, a transcriptional repressor and one of the main downstream effectors of Notch signaling. At best of our knowledge, this is also the first biophysical investigation on a DNA-binding protein with a bHLH-ORANGE domain architecture. From our data, we propose a functional model of HES-1 that, albeit speculative, explains how different degrees of order/disorder in the basic, helix-loop-helix, ORANGE, and C-terminal domains may govern not only binding to DNA but also protein dimerization.

Acknowledgements

We are in debt to Dr. Doriano Lamba (CNR-ELETTRA, Trieste, Italy) for the extensive use of the CD spectropolarimeter. We are grateful to Dr. Matija Popovic (ICGEB, Trieste) for the expression and purification of H1CT. We also thank Prof. Ilian Jelezarov (University of Zürich, Switzerland) for helpful suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.bbapap.2010.08.010.

References

[1] A. Fischer, M. Gessler, Delta-Notch—and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors, Nucleic Acids Res. 35 (2007) 4583–4596.
[2] R. Kageyama, T. Ohtsuka, T. Kobayashi, The Hes gene family: repressors and oscillators that orchestrate embryogenesis, Development 134 (2007) 1243–1251.
[3] T. Iso, L. Kedes, Y. Hamamori, HES and HERP families: multiple effectors of the Notch signaling pathway, J. Cell. Physiol. 194 (2003) 237–255.
[4] R.L. Davis, D.L. Turner, Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning, Oncogene 20 (2001) 8342–8357.
[5] R. Kageyama, T. Ohtsuka, K. Tomita, The bHLH gene Hes1 regulates differentiation of multiple cell types, Mol. Cells 10 (2000) 1–7.
[6] H. Hirata, S. Yoshiura, T. Ohtsuka, Y. Bessho, T. Harada, K. Yoshikawa, R. Kageyama, Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop, Science 298 (2002) 840–843.
[7] S. Jarriault, C. Brou, F. Logeat, E.H. Schroeter, R. Kopan, A. Israel, Signalling downstream of activated mammalian Notch, Nature 377 (1995) 355–358.
[8] L. Sang, H.A. Coller, J.M. Roberts, Control of the reversibility of cellular quiescence by the transcriptional repressor HES1, Science 321 (2008) 1095–1100.
[9] K. Takebayashi, Y. Sasai, Y. Sakai, T. Watanabe, S. Nakanishi, R. Kageyama, Structure, chromosomal locus, and promoter analysis of the gene encoding the mouse helix-loop-helix factor HES-1. Negative autoregulation through the multiple N box elements, J. Biol. Chem. 269 (1994) 5150–5156.
[10] H. Chen, A. Thiagalingam, H. Chopra, M.W. Borges, J.N. Feder, B.D. Nelkin, S.B. Baylin, D.W. Ball, 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 94 (1997) 5355–5360.
[11] P. Castella, S. Sawai, K. Nakao, J.A. Wagner, M. Caudy, HES-1 repression of differentiation and proliferation in PC12 cells: role for the helix 3–helix 4 domain in transcription repression, Mol. Cell. Biol. 20 (2000) 6170–6183.
[12] K. Murata, M. Hattori, N. Hirai, Y. Shinozuka, H. Hirata, R. Kageyama, T. Sakai, N. Minato, Hes1 directly controls cell proliferation through the transcriptional repression of p27Kip1, Mol. Cell. Biol. 25 (2005) 4262–4271.
[13] Y. Sasai, R. Kageyama, Y. Tagawa, R. Shigemoto, S. Nakanishi, Two mammalian helix-loop-helix factors structurally related to Drosophila hairy and Enhancer of split, Genes Dev. 6 (1992) 2620–2634.
[14] H. Hirata, T. Ohtsuka, Y. Bessho, R. Kageyama, Generation of structurally and functionally distinct factors from the basic helix-loop-helix gene Hes3 by alternative first exons, J. Biol. Chem. 275 (2000) 19083–19089.
[15] S. Bae, Y. Bessho, M. Hojo, R. Kageyama, The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation, Development 127 (2000) 2933–2943.
[16] O. Nakagawa, D.G. McFadden, M. Nakagawa, H. Yanagisawa, T. Hu, D. Srivastava, E.N. Olson, Members of the HRT family of basic helix-loop-helix proteins act as transcriptional repressors downstream of Notch signaling, Proc. Natl Acad. Sci. USA 97 (2000) 13655–13660.
[17] T. Iso, V. Sartorelli, C. Poizat, S. Iezzi, H.Y. Wu, G. Chung, L. Kedes, Y. Hamamori, HERP, a novel heterodimer partner of HES/E(spl) in Notch signaling, Mol. Cell. Biol. 21 (2001) 6080–6089.
[18] D. Grbavec, S. Stifani, Molecular interaction between TLE1 and the carboxyl-terminal domain of HES-1 containing the WRPW motif, Biochem. Biophys. Res. Commun. 223 (1996) 701–705.
[19] L. Whitmore, B.A. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, Nucleic Acids Res. 32 (2004) W668–W673.

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010

[20] R.L. Winston, D.P. Millar, J.M. Gottesfeld, S.B. Kent, Characterization of the DNA binding properties of the bHLH domain of Deadpan to single and tandem sites, Biochemistry 38 (1999) 5138–5146.
[21] D. Meierhan, C. el-Ariss, M. Neuenschwander, M. Sieber, J.F. Stackhouse, R.K. Allemann, DNA binding specificity of the basic-helix-loop-helix protein MASH-1, Biochemistry 34 (1995) 11026–11036.
[22] A.G. Kunne, D. Meierhans, R.K. Allemann, Basic helix-loop-helix protein MyoD displays modest DNA binding specificity, FEBS Lett. 391 (1996) 79–83.
[23] R.L. Winston, J.A. Ehley, E.E. Baird, P.B. Dervan, J.M. Gottesfeld, Asymmetric DNA binding by a homodimeric bHLH protein, Biochemistry 39 (2000) 9092–9098.
[24] C.S. Tremblay, F.F. Huang, O. Habi, C.C. Huard, C. Godin, G. Levesque, M. Carreau, HES1 is a novel interactor of the Fanconi anemia core complex, Blood 112 (2008) 2062–2070.
[25] A. Strom, P. Castella, J. Rockwood, J. Wagner, M. Caudy, Mediation of NGF signaling by post-translational inhibition of HES-1, a basic helix-loop-helix repressor of neuronal differentiation, Genes Dev. 11 (1997) 3168–3181.
[26] B.G. Ju, D. Solum, E.J. Song, K.J. Lee, D.W. Rose, C.K. Glass, M.G. Rosenfeld, Activating the PARP-1 sensor component of the groucho/TLE1 corepressor complex mediates a CaMKinase IIdelta-dependent neurogenic gene activation pathway, Cell 119 (2004) 815–829.

Please cite this article as: M. Coglievina, et al., Different degrees of structural order in distinct regions of the transcriptional repressor HES-1, Biochim. Biophys. Acta (2010), doi:10.1016/j.bbapap.2010.08.010