Identification of a region on hypoxia-inducible-factor prolyl 4-hydroxylases that determines their specificity for the oxygen degradation domains

Diego VILLAR*, Alicia VARA-VEGA†, Manuel O. LANDÁZURI†¹ and Luis DEL PESO*^{1,2}

*Departamento de Bioquímica, Instituto de Investigaciones Biomédicas 'Alberto Sols', Consejo Superior de Investigaciones Científicas–Universidad Autónoma de Madrid, Arturo Duperier 4, 28029 Madrid, Spain, and †Servicio de Inmunología, Hospital de la Princesa–Universidad Autónoma de Madrid, Diego de León 62, 28006 Madrid, Spain

HIFs [hypoxia-inducible (transcription) factors] are essential for the induction of an adaptive gene expression programme under low oxygen partial pressure. The activity of these transcription factors is mainly determined by the stability of the HIF α subunit, which is regulated, in an oxygen-dependent manner, by a family of three prolyl 4-hydroxylases [EGLN1–EGLN3 (EGL nine homologues 1–3)]. HIF α contains two, N- and C-terminal, independent ODDs (oxygen-dependent degradation domains), namely NODD and CODD, that, upon hydroxylation by the EGLNs, target HIF α for proteasomal degradation. *In vitro* studies indicate that each EGLN shows a differential preference for ODDs. However, the sequence determinants for such specificity are unknown. In the present study we showed that whereas EGLN1 and EGLN2 acted upon any of these ODDs to regulate HIF1 α protein levels and activity *in vivo*, EGLN3 only acted on the CODD. With the aim of identifying the region within

EGLNs responsible for their differential substrate preference, we investigated the activity and binding pattern of different EGLN deletions and chimaeric constructs generated by domain swapping between EGLN1 and EGLN3. These studies revealed a region of 97 residues that was sufficient to confer the characteristic substrate binding observed for each EGLN. Within this region, we identified the minimal sequence (EGLN1 residues 236–252) involved in substrate discrimination. Importantly, mapping of these sequences on the EGLN1 tertiary structure indicates that substrate specificity is determined by a region relatively remote from the catalytic site.

Key words: EGL nine homologue (EGLN), hypoxia, hypoxia-inducible factor (HIF), oxygen-degradation domain (ODD), proline hydroxylase domain (PHD), von Hippel–Lindau protein (VHL protein).

INTRODUCTION

Cells respond to a decreased oxygen partial pressure with the induction of an adaptive gene expression programme aimed at restoring oxygen supply and maintaining cell viability during oxygen restriction. Most of the genes induced during low oxygen exposure are under control of the HIF (hypoxia-inducible factor) transcription factors [1]. HIFs are heterodimers of a constitutively expressed β -subunit [HIF β , also known as ARNT (aryl hydrocarbon receptor nuclear translocator)] and one of the three known oxygen-regulated α -subunits (HIF1 α , HIF2 α or HIF3 α). Both α - and β -subunits are basic helix–loop–helix transcription factors containing a PAS [PER–AhR–SIM (period gene protein–aryl hydrocarbon receptor–single-minded protein)] domain involved in protein–protein interaction. HIF α subunits are extremely short-lived under normal oxygen partial pressure, owing to efficient proteasomal degradation [2]. HIF α degradation is triggered by the hydroxylation of specific proline residues [3,4]. This reaction is catalysed by a family of proteins homologous with the product of the *Caenorhabditis elegans* *egl-9* gene [5–7] and are thus named EGLNs (EGL nine homologues) 1, 2 and 3 [8]. These enzymes are also widely known as PHDs (proline hydroxylase domains) 2, 1 and 3 respectively. Hydroxylated HIF α subunits are specifically recognized and targeted for degradation by an E3-ubiquitin ligase complex containing the VHL (von Hippel–Lindau) protein [9]. EGLNs require molecular oxygen for the hydroxylation reaction. Since the K_m values of these enzymes for oxygen are high, their

activity changes within a physiologically relevant range of oxygen concentrations and they are hence thought to function as molecular oxygen sensors in the HIF regulation pathway [10,11]. Thus, under limiting oxygen partial pressure, proline hydroxylation is compromised and HIF α escapes VHL protein binding and degradation. As a result, the half-life of HIF α is extended, resulting in its accumulation and interaction with HIF β , forming a transcriptionally competent dimer [2].

Two independent proline residues within HIF1 α , namely P402 and P564, have been found to be hydroxylated in an oxygen-dependent fashion and have been proposed to be substrates for the EGLNs [12]. Interestingly, the sequences containing both proline residues, as well as the equivalent proline residues in HIF2 α (P405 and P531) and HIF3 α (P491), align into a conserved LXXLAP motif [3,4]. Moreover, this sequence motif is evolutionarily conserved in HIF α proteins from nematodes (e.g. *Caenorhabditis*) to mammals, suggesting an important role of these residues for EGLN and/or VHL binding [3,4]. However, several studies suggested the absence of a rigid consensus sequence for EGLN binding [13,14]. In spite of this apparent flexibility in their binding to target sequences, different EGLNs show a differential preference for the sequence containing HIF1 α P402 [NODD (N-terminal ODD (oxygen-dependent degradation domain))] or HIF1 α P564 [CODD (C-terminal ODD)] residues [10,11,14–16]. Both NODD and CODD are substrates *in vitro* for EGLN1 and EGLN2, although the K_m value for NODD is much lower [11]. By contrast, EGLN3 only hydroxylates CODD [10,11,14]. In addition to these *in vitro* studies, it has been demonstrated that

Abbreviations used: AD, Gal4 activation domain; DBD, Gal4 DNA-binding domain; EGLN, EGL nine homologue; HIF, hypoxia-inducible factor; HRE, hypoxia-response element; MYND, myeloid, Nervy, and DEAF-1; NODD, N-terminal ODD (oxygen-dependent degradation domain); CODD, C-terminal ODD; SD, synthetic defined; VHL, von Hippel–Lindau.

¹ These authors contributed equally to this work.

² To whom correspondence should be addressed (email lpeso@iib.uam.es).

both isolated ODDs are sufficient to drive protein degradation in mammalian cells [15]. However, the relative contribution of each site in the regulation of full-length HIF α or the specificity of each EGLN for the different sites *in vivo* are issues that await full elucidation. Therefore, although EGLNs have different substrate preferences, both *in vivo* and *in vitro*, the sequence determinants responsible for this substrate discrimination remain unclear. Moreover, in contrast with the effort dedicated to the identification of the sequence determinants in substrates, to the best of our knowledge no attention has been paid to the regions within EGLNs responsible for their different target preference.

In the present study we investigated the regulation of each ODD within full-length HIF1 α by individual EGLNs *in vivo*. We found that, in agreement with *in vitro* data, EGLN1 and EGLN2 were able to regulate HIF by acting on either of the two ODDs, whereas EGLN3 preferentially recognized CODD. Then we investigated which region of these enzymes was responsible for the observed substrate-preference pattern. Through the analysis of the effect of EGLN deletions and chimaeric constructs on HIF activity and stability, together with the utilization of yeast binding assays, we identified a region that conferred the specific EGLN substrate preference. Interestingly, this region is located relatively far from the catalytic site where proline hydroxylation occurs.

MATERIALS AND METHODS

Cells and reagents

For the yeast two-hybrid assays described in the present paper we used the reporter yeast strain AH109 from BD Biosciences (Palo Alto, CA, U.S.A.). Yeast growth-media reagents [YPD (yeast extract/peptone/dextrose), SD (synthetic defined) amino acids and X-gal (5-bromo-4-chloroindol-3-yl β -D-galactopyranoside)] were from BD Biosciences. Reporter assays were performed using the human cervical-carcinoma cell line HeLa. HeLa cells were maintained in Dulbecco's modified Eagle's medium (Cambrex Bio Science, Walkersville, MD, U.S.A.), supplemented with 100 units/ml penicillin, 100 μ g/ml streptomycin and 5% (v/v) fetal bovine serum. Cells were grown at 37°C and 5% CO₂ in a humidified incubator.

Anti-HIF1 α monoclonal antibody was purchased from Transduction Laboratories (Lexington, KY, U.S.A.), whereas anti- α -tubulin and anti-[FLAG epitope (M2 clone)] antibodies were from Sigma (St Louis, MO, U.S.A.).

Generation of plasmid constructs

EGLNs cDNAs were generously provided by Dr Steven L. McKnight (Graduate School of Biomedical Sciences, University of Texas Southwestern Medical Center at Dallas, Dallas, TX, U.S.A.) [6] and Professor Peter Ratcliffe (Henry Wellcome Building for Molecular Physiology, University of Oxford, Oxford, U.K.) [4]. Chimaeric EGLNs were generated by PCR, and the resulting coding sequences were cloned into the pCDNA-FLAG plasmid for their expression in mammalian cells or into the pBRIDGE plasmid for the yeast two-hybrid assays. The identity of all constructs was confirmed by sequencing.

Constructs expressing wild-type and mutant forms of HIF1 α (P402A, P564A and P402A/P564A) have been described previously [16].

Yeast transformation and interaction assays

Freshly prepared competent yeast cells were transformed with 0.1 μ g of each of the indicated plasmids by a modified version of the lithium acetate method [17]. The semi-quantitative interaction

experiments have been described previously [16]. Briefly, transformed yeast were plated on minimal SD media plates in the absence of leucine and tryptophan. Subsequently, an equal number of colonies from each transformation were transferred to 0.9% NaCl and subjected to serial dilutions. Aliquots of each cell suspension (typically 20 μ l) were plated on culture media with different stringencies: (1) plates lacking leucine/tryptophan (minimal stringency; no interaction between fusion proteins required for yeast growth); (2) plates lacking leucine, tryptophan and histidine (medium stringency; interaction required to support yeast growth); or (3) plates lacking leucine, tryptophan, histidine and adenine (maximal stringency; interaction required to support yeast growth).

Cell transfection and reporter assays

Cells were seeded on six-well (10⁵ cells/well) or 12-well plates (4 $\times$ 10⁴ cells/well) 16 h prior to transfection. A 9 μ g DNA mixture containing 1 μ g of p9 $\times$ HRE-Luc reporter plasmid [18], 2.5 μ g of a construct directing the expression of HIF1 α wild-type, HIF1 α P402A, HIF1 α P564A or HIF1 α P402A/P564A, and 50 ng of a plasmid encoding for *Renilla* (sea pansy) luciferase under the control of a simian-virus-40 promoter was used for transfection using the calcium phosphate method [19]. The p9 $\times$ HRE-Luc plasmid contains the firefly luciferase coding sequence under the control of nine consecutive copies of the vascular-endothelial-growth-factor HRE (hypoxia response element) [18]. Where indicated, an additional 200 ng of a pCDNA-FLAG-EGLN construct was added to the mixture. At 6 h after transfection, cells were washed and further incubated for 14 h.

For reporter assays, cells in 12-well plates were lysed and the firefly and *Renilla* luciferase activities of the lysate were determined using a dual-luciferase system (Promega, Madison, WI, U.S.A.). The firefly luciferase activity was normalized to that of *Renilla* luciferase. The results are presented as the average reporter activity obtained in *n* (indicated in each Figure legend) independent experiments as the percentage of the activity obtained in cells expressing HIF1 α protein in the absence of exogenous EGLN enzymes ('none'). The luciferase activity observed in cells transfected with the reporter construct alone is shown ('Bkgr'). The error bars in graphs showing reporter activity represent the S.E.M.

Western blot

Cells were washed with ice-cold TBS (Tris-buffered saline; 50 mM Tris, pH 7.6, and 150 mM NaCl) and harvested in 80 μ l of NP40 lysis buffer [1% Nonidet P40, 140 mM NaCl, 10 mM EDTA, 10% (v/v) glycerol, 20 mM Tris, pH 8, 1 mM PMSF and 1 mM orthovanadate]. Samples (30 μ g of total protein) were resolved on SDS/10%-(w/v)-polyacrylamide gels. Proteins were then transferred on to nitrocellulose membranes (Bio-Rad, Hercules, CA, U.S.A.), blocked with 5% non-fat dry milk in TBS-T (TBS + 0.1% Tween 20) and incubated overnight at 4°C with the indicated antibodies. Immunolabelling was detected by ECL[®] (enhanced chemiluminescence; Amersham Pharmacia Biotech, Piscataway, NJ, U.S.A.) and visualized with a digital luminescent image analyser (LAS-1000 CH; Fujifilm, Tokyo, Japan).

Model of EGLN3 structure and generation of structure images

A model of EGLN3 structure was generated by homology modelling using EGLN1 structure (PDB 2619) as a template. The model was generated on the CPHmodels-2.0 Server at the Technical University of Denmark, Kgs. Lyngby, Denmark (<http://www.cbs.dtu.dk/services/CPHmodels/>). A similar model of EGLN3 structure was obtained using the SwissModel Server (Biozentrum,

University of Basel, Basel, Switzerland, and the Advanced Biomedical Computing Center, National Cancer Institute, Frederick, MD, U.S.A.; <http://swissmodel.expasy.org/>). PyMOL software (DeLano Scientific LLC, Palo Alto, CA, U.S.A.; <http://www.pymol.org>) was used to visualize the EGLN1 structure (Protein Data Bank ID 2G19) and the EGLN3 model and to generate all the images in Figure 6 (below), including the calculation of the electrostatic potentials.

Statistical analysis of data

Experimental data were analysed with the Prism™ GraphPad (version 4.01) software. Data from reporter assays were analysed by ANOVA, followed by the Bonferroni's multiple comparison test. For the multiple comparison the activity of each chimaeric construct was compared with that of wild-type EGLN3 or EGLN1 (Figure 2 below). Significant differences ($P < 0.001$) are indicated in Figures by asterisks. Since the transcriptional activity of HIF in the absence of hydroxylases was normalized, these values were not included in the statistical analysis.

RESULTS

Individual EGLN paralogues show a differential preference for each HIF ODD

In vitro enzymic studies have demonstrated that, unlike EGLN1 and EGLN2, EGLN3 is unable to hydroxylate the NODD, even in the context of the full-length HIF1 α protein [10,11]. These results are in agreement with the finding that EGLN1 and EGLN2, but not EGLN3, are able to target isolated NODD in mammalian cells [15]. However, others have found hydroxylation of P402 upon expression of EGLN3 in mammalian cells [20]. In order to test the substrate specificities of the EGLNs *in vivo* in the context of the full-length HIF α protein, we studied the effect of EGLNs overexpression on the transcriptional activity and stability of different HIF α mutant proteins (Figure 1).

Overexpression of wild-type HIF α resulted in the saturation of the EGLN/VHL degradation pathway, leading to HIF α activity as a consequence of protein accumulation (Figure 1A). Co-expression of any of the EGLN paralogues, together with wild-type HIF α , overcame the pathway blockage, resulting in a large decrease on HIF activity and protein levels (Figure 1A). By contrast, the levels and activity of a HIF1 α protein containing the P402A and P564A mutations was resistant to the expression of EGLNs (Figure 1D). Thus, and in accordance with the biochemistry of prolyl hydroxylases, the effect of EGLNs on HIF1 α activity and stability is dependent on the presence of the target proline residues. Next we investigated the individual contribution of each ODD to the regulation of HIF by each EGLN. Transfection of HIF1 α -P402A (Figure 1B) or HIF1 α -P564A (Figure 1C) single mutants resulted in HIF transcriptional activity and protein accumulation. Importantly, co-expression of EGLN1 or EGLN2 down-regulated protein levels and suppressed the transcriptional activity induced by either of the mutant forms (Figures 1B and 1C). In sharp contrast, EGLN3 overexpression was able to repress HIF1 α -P402A (Figure 1B), containing an intact CODD, but had no detectable effect on HIF1 α -P564A protein accumulation or its activity (Figure 1C).

In vitro studies [10,11] showed that EGLN3 has a very high K_m value for HIF1 α NODD that correlates with the low activity of EGLN3 towards HIF1 α -P564A observed in our assays. A high K_m value might reflect a reduced binding affinity of EGLN3 for NODD, but only when the dissociation constant for the enzyme-substrate complex is larger than k_{cat} . Thus we next studied whether

differential binding of the EGLN enzymes to the NODD could explain the observed activity. In order to avoid interference with the cellular endogenous regulatory machinery and determine the interaction strength semi-quantitatively, we employed the yeast two-hybrid system to test these interactions [16]. To this end, EGLN1 and EGLN3 coding sequences were cloned in-frame with the activation domain of the yeast Gal4 transcription factor (termed AD). The derived constructs were transformed into yeast together with a plasmid encoding for full-length HIF1 α as a fusion protein with the DNA binding domain from Gal4 transcription factor (termed DBD). Binding between these fusion proteins allows the growth of yeast in the absence of specific nutrients. As shown in Figure 1(E), the three EGLN-Gal4AD fusion proteins strongly bound to full-length wild-type HIF1 α , as demonstrated by the growth of yeast cells in the conditions of highest stringency ('INTERACTION', Figure 1E). As expected, no yeast growth was observed under stringent conditions when the double mutant HIF1 α -P402A/P564A protein was tested against any of the EGLNs (Figure 1E). This finding indicated that binding was completely dependent on the presence of the target proline residue. Importantly, whereas EGLN1 and EGLN2 were able to bind both HIF1 α -P402A and HIF1 α -P564A single mutants, EGLN3 only bound the HIF1 α -P402A mutant (Figure 1E). This result strongly suggests that the lack of EGLN3 activity toward the NODD observed in mammalian cells (Figure 1C) is due to decreased binding to the P402 region.

The N-terminal unique region of EGLN1 does not determine substrate specificity

Since EGLN1/2 and EGLN3 show different substrate preferences, we reasoned that they might have different substrate-binding sites. Alignment of the three EGLNs paralogues (Figure 2) showed that they shared a highly conserved C-terminal region. In particular, 87 % residue identity was observed between EGLN1 and EGLN3 within this region (labelled C1 and C3 respectively, Figure 2). This region contains the residues involved in iron co-ordination and 2-oxoglutarate binding (Figure 2, arrowheads) and is therefore thought to form the catalytic site [21]. This sequence is preceded by a region that shares a more modest conservation between family members (45 % residue identity between EGLN1 and EGLN3 within the regions labelled as N1 and N3 respectively). A similar distribution of conserved residues within these regions was observed when EGLN1 was aligned with EGLN2 (Figure 2). Finally, EGLN1 contains a unique N-terminal region (labelled U1 in Figure 2) that is completely missing in EGLN3 and shows no detectable conservation with the equivalent region in EGLN2.

Thus we focused on these regions of differential conservation in order to identify the region determining the substrate specificity of prolyl hydroxylases. Since the unique N-terminal region (marked U in Figure 2) is only present in the EGLN isoforms that recognized NODD (EGLN1 and EGLN2), we firstly addressed the role of this region on substrate selection. To this end, we compared the substrate specificity of wild-type EGLN1 with that of a construct lacking the unique N-terminal region comprising residues 1–177 (Figure 3A, Δ U1). Additionally, we investigated the activity of an EGLN1 PCR variant (Figure 3A, Δ EGLN1) that lacks 100 residues in its N-terminal region, from Gly⁷⁶ to Gly¹⁷⁷, but conserves the MYND (myeloid, Nervy, and DEAF-1)-type zinc finger (residues 21–58). As shown in Figure 3, neither deletion of amino acids Gly⁷⁶–Gly¹⁷⁷ nor the whole unique N-terminal region (residues 1–177) had an effect on substrate selection by EGLN1. Both EGLN1 truncated forms were able to suppress both activity and protein levels of all HIF1 α forms, except for the double mutant (Figures 3B–3E).

Figure 1 Effect of EGLN expression on HIF1 α protein levels and activity

(A) HeLa cells were transfected with an HRE-driven reporter construct together with a plasmid encoding for HIF1 α (none) or a combination of plasmids encoding for HIF1 α and the indicated EGLN enzymes. The graphs show the normalized average reporter activity obtained in three independent experiments ($n = 3$). The HIF1 α and α -tubulin (loading control) protein levels were determined in HeLa cell cultures transfected with the indicated combination of plasmids detailed above. (B–D) Cells were treated as in (A), except that P402A (B), P564A (C) or P402A/P564A (D) mutant forms of HIF1 α were used for the assays. The asterisks (***) indicate that the mean values for EGLN1 and EGLN2 are significantly different ($P < 0.001$) from those obtained for EGLN3. (E) Yeast cells were transformed with AD–HIF1 α constructs together with DBD–EGLN1, DBD–EGLN2 or DBD–EGLN3 proteins. Serial dilutions of transformed clones were grown on plates lacking leucine and tryptophan (CONTROL, no interaction required between fusion proteins for yeast growth) or lacking leucine, tryptophan, histidine and adenine (INTERACTION, the interaction between fusion proteins is required for yeast growth). WT, full-length wild-type HIF1 α ; P402A, HIF1 α P402A mutant (intact CDD); P564A, HIF1 α P564A mutant (intact NODD); PP, HIF1 α P402A/P564A mutant. The results shown are representative for at least three independent experiments. Bkgr, luciferase activity observed in cells transfected with the reporter construct alone; Tub., α -tubulin; WB, Western blot.

Accordingly, these truncated EGLN1 forms bound wild-type and single-mutant HIF1 α proteins similarly to full-length EGLN1 protein (Figure 3F). Similarly to EGLN1, deletion of the N-terminal unique region (residues 1–167) from EGLN2 had no effect on its activity on both ODDs (results not shown). Thus the results shown in Figure 3 indicate that the N-terminal unique region is not required to determine EGLN substrate preference.

Identification of an EGLN region that determines their differential substrate specificity

In order to determine the region responsible for the differential substrate specificity of each EGLN while retaining the domain architecture of the wild-type enzymes, chimaeric constructs were generated by swapping the regions identified by sequence

Figure 2 Sequence conservation among EGLNs paralogues

Human EGLN1 (NP_071344) was pairwise-aligned with EGLN3 (NP_071356) or EGLN2 (NP_444274) protein sequences using the BLAST algorithm [29]. The graph represents the percentage of sequence identity within a 20 residue sliding window. Positions correspond to EGLN1 sequence. Diagrams above the graph represent the aligned EGLN sequences. Boxes correspond to regions of none (U1/U2), moderate (N1–N3) or high (C1–C3) sequence conservation that were swapped for the generation of chimaeras. Black arrowheads indicate the location of the conserved triad His-Asp-His of residues involved in iron binding at the catalytic centre. Numbers indicate amino acid residue positions.

comparison (Figure 2) between EGLN1 and EGLN3 (Figure 4A). Substitution of the N-terminal region of EGLN3 (residues 1–118, labelled N3 in Figure 2) by that of EGLN1 (residues 1–296) resulted in an enzyme (U1N1C3) with the ability to act upon NODD as well as CODD, in a pattern identical with that of EGLN1 (Figures 4B–4E). Moreover, transference of the EGLN1 unique N-terminal region (U1) to EGLN3 did not alter its substrate preference, since the resulting enzyme (U1N3C3) behaved as wild-type EGLN3 (Figure 4). Hence the EGLN1 residues 201–296 (N1 region) were sufficient to confer the ability to act upon both ODDs, while residues 1–200 (U1 region) were not. Conversely, transference of the N-terminal region from EGLN3 (residues 1–118, labelled N3) upstream of the catalytic domain of EGLN1 produced an enzyme (N3C1) that suppressed wild-type HIF1 α (Figure 4B) and the HIF1 α -P402A (Figure 4C) mutant. However, N3C1 expression did not decrease the level nor the activity of the HIF1 α -P564A mutant protein (Figure 4D), in a pattern identical with that of full-length EGLN3. Furthermore, analysis of the binding pattern of these chimaeric constructs (Figure 4F) showed that enzymes harbouring residues 201–296 from EGLN1 (U1N1C3 and N1C3; Figure 4E and results not shown) bound to both NODD and CODD, whereas constructs containing the equivalent region from EGLN3 (U1N3C3 and N3C1; Figure 4F) failed to bind NODD with detectable strength, a finding in perfect agreement with the observed activity towards ODDs (Figures 4B–4E). Altogether, these results identified a region responsible for the determination of EGLN substrate-binding specificity, which comprises residues 201–296 in EGLN1, and its homologue in EGLN3.

We exploited the recently described crystal structure of EGLN1 (residues 188–403) [21] to locate the N-region (residues 201–298) on the folded structure (Figure 5). The opening to the catalytic site lies in one margin of the surface of the C-region (Figures 5A

and 5B, coloured orange), very close to the boundary with the N-region (Figures 5A, coloured in blue tones). Thus a large fraction of the N-region surface will be accessible to the substrate approaching the catalytic site (Figures 5A and 5B). Structural differences might underlie the variation in substrate selection displayed by EGLN1 compared with EGLN3. Consequently, and in order to approach the comparison of the two enzymes and further delimit the EGLN sequence responsible for substrate specificity, we generated a model of EGLN3 tertiary structure by homology modelling [22] using EGLN1 as the template (Figures 5C and 5D). This model suggests that the overall structure of EGLN3 is very similar to that of EGLN1 (Figure 5A). However, owing to their residue dissimilarities (Figure 5E), the surface electrostatic potential of these molecules is markedly different at specific patches (compare Figures 5C and 5D). The most prominent difference maps on the finger-like projection within the N-region (Figures 5C and 5D). Although the surface of this structure in the EGLN1 molecule presents alternate acidic and basic regions (Figure 5D), that of EGLN3 was predicted to be uniformly charged (Figure 5C). To test the role of this finger-like structure on substrate selection, we generated two more chimaeric constructs (Figure 6A), in which the N-region was composed of sequences from EGLN1 (N1) and EGLN3 (N3) in different proportions. Importantly, N1/3A and N1/3B constructs differ only in 16 residues (residues 236–252), corresponding to the sequence of the finger-like structure.

Analysis of the activity of these constructs (Figures 6B–6E), showed that both N1/3A and N1/3B were able to inhibit the wild-type and the P402A mutant. However, only N1/3A had a significant effect on the P564A mutant ($P < 0.01$, N1/3A compared with U1N3C3), whereas the effect of N1/3B was undistinguishable from U1N3C3 ($P > 0.05$, N1/3B compared with U1N3C3). These effects were even clearer at the level of binding to substrate. As shown in Figure 6(F), whereas N1/3B only bound CODD within HIF1 α , the construct N1/3A recognized CODD as well as NODD, in a binding pattern identical with that of U1N1C3 and EGLN1. Thus these results indicate that the first half of the N1 region (EGLN1 residues 201–252) is sufficient to confer the ability to bind P402 and that the residues 236–252 (shown in dark blue in Figures 5A and 5B) are necessary. Additionally, our results strongly suggest that the finger-like structure plays an important role in substrate binding. In contrast, residues 253–298 (shown in blue in Figures 5A and 5B) did not seem to play a role in binding NODD, regardless of their localization surrounding the entrance to the catalytic site.

DISCUSSION

Several studies have indicated that EGLNs have different specificities for isolated NODD and CODD sequences [10,15,16,20]. In the present study, we investigated the different activities of individual EGLNs toward each ODD within the context of the full-length HIF1 α protein. In agreement with previous findings, in which isolated ODDs were used [15], we found that EGLN1 and EGLN2 work on both NODD and CODD, whereas EGLN3 only has significant activity upon CODD. We also demonstrated that the effect of individual EGLNs on each ODD correlates with their ability to bind them, strongly suggesting that the differential activity observed is a result of differences in binding affinity rather than differences on catalytic activity toward each target sequence. Our present results, and those of others [15], demonstrate that any of the sites is sufficient to regulate HIF α levels and activity. However, we found that expression of the P402A mutant resulted in lower HIF α levels

Figure 3 Effect of different N-terminal deletions on EGLN1 activity

(A) Schematic diagram representing the constructs generated by deletion of the indicated regions of EGLN1. (B) HeLa cells were transfected with an HRE-driven reporter construct, together with a plasmid encoding for HIF1 α (none) or a combination of plasmids encoding for HIF1 α and the indicated EGLN1 forms. Δ U1, truncated EGLN1 lacking residues 1–177. Δ EGLN1, truncated EGLN1 lacking residues 76–177. The results are normalized average reporter activities obtained in three independent experiments ($n = 3$). The HIF1 α and α -tubulin (loading control) protein levels were determined in HeLa cell cultures transfected with the indicated combination of plasmid as detailed above. (C–E) Cells were treated as in (A), except that P402A (C), P564A (D) or P402A/P564A (E) mutant forms of HIF1 α were used for the assays. (F) Yeast cells were transformed with AD–HIF1 α constructs together with wild-type DBD–EGLN1 or the indicated deletion mutants fused to the DBD. Serial dilutions of transformed clones were grown on plates lacking leucine and tryptophan (CONTROL) or plates lacking leucine, tryptophan and histidine (INTERACTION). The results shown are representative for at least three independent experiments. Abbreviations and symbols are as described for Figure 1.

and activity than did expression of the P564A mutant (results not shown). This probably reflects the combined effects of the lower activity of EGLN1 and EGLN2 on NODD compared with CODD [10,11] and the lack of effect of endogenous EGLN3 on NODD.

Interestingly, we observed that the ability of EGLNs to down-regulate HIF activity was consistently stronger than their effect on HIF1 α protein levels. This was particularly apparent in the case of EGLN2 (Figures 1A–1C). Several recent reports [23–25] suggest that EGLNs might have an effect on HIF activity independent of

Figure 4 Effect of chimaeric EGLN constructs on HIF1 α protein levels and activity

(A) Schematic diagram representing the chimaeric constructs generated by exchange of the indicated regions between EGLN1 and EGLN3. (B) HeLa cells were transfected with an HRE-driven reporter construct together with a plasmid encoding for HIF1 α (none) or a combination of plasmids encoding for HIF1 α and the indicated wild-type or chimaeric enzymes. A description of the domain structure of the chimerical constructs structure is given in Figure 2. The graphs represent the normalized average reporter activity obtained in three independent experiments ($n = 3$). The asterisks (***) indicate that mean values for those constructs are significantly different ($P < 0.001$) from those obtained for EGLN3. The HIF1 α and α -tubulin (loading control) protein levels were determined in HeLa cell cultures transfected with the indicated combination of plasmids as detailed above. (C–E) Cells were treated as in (A), except that P402A (C), P564A (D) or P402A/P564A (E) mutant forms of HIF1 α were used for the assays. (F) Yeast cells were transformed with AD–HIF1 α constructs together with DBD–EGLN1 or DBD–EGLN3 proteins. Serial dilutions of transformed clones were grown on plates lacking leucine and tryptophan (CONTROL) or plates lacking leucine, tryptophan, histidine and adenine (INTERACTION). The results shown are representative for at least three independent experiments. The abbreviations and symbols are as described for Figure 1.

Figure 5 Localization and electrostatic potential of the N-region on EGLN1 structure and EGLN3 model

(A,B and D) Images of the molecular surface of EGLN1 (Protein Data Bank 2G19). The colour code in (A) and (B) was as follows: cyan, residues 188–236; dark blue, residues 237–252; light blue, residues 253–298; and orange, residues 297–403. The wire representation of the 2-oxoglutarate competitive inhibitor compound A is shown in red, and the iron atom is represented as a yellow sphere. (A) Front view of the molecule showing the deep catalytic site; (B) 180° rotation of the molecule depicted in (A); (C and D) representation of the qualitative electrostatic surface potential (PyMol) for the EGLN1 structure (D) and a EGLN3 model (C). Shades of red and blue are used to represent negative and positive potentials respectively; (E) alignment of human EGLN1 and EGLN3 sequences. The numbers correspond to residue positions according to the EGLN1 sequence. Identical residues are shown on a black background and conservative substitutions have a grey background.

their ability to regulate HIF α stability, suggesting an explanation for our observation.

Analysis of the substrate binding and activity of the different EGLN constructs toward HIF α mutants led us to identify a region (corresponding to residues 201–296 in the EGLN1 sequence) responsible for the differential target preference of EGLNs. This region showed only a moderate conservation among family members, in contrast with the high conservation of the C-terminal region containing the catalytic residues of EGLN enzymes. The variation observed in this region might be required to accommodate the different substrate preferences of these enzymes. Consistently, alignment of the human EGLN paralogues revealed that, within the region responsible for substrate discrimination (EGLN1 residues 201–236), EGLN1 has higher sequence similarity with EGLN2 (56.3% identity) than with EGLN3 (44.0% identity) (Figure 2). This is not a mere consequence of closer evolutionary distance between EGLN1

and EGLN2, since, within the catalytic region (corresponding to residues 299–426 in the EGLN1 protein), the sequence identity between these two proteins is slightly lower (81.9%) than between EGLN1 and EGLN3 (86.0%) (Figure 2). The lack of detectable conservation between EGLN1 and EGLN2 within their first 200 residues, together with our observation that both enzymes shared a common substrate specificity (Figure 1), argues against a role for this region in substrate selection, a conclusion that is also supported by our results (Figure 3). One possibility is that this region regulates EGLN1 in response to intracellular cues. In agreement with this, a recent report [26] suggested that the unique N-terminal region of EGLN1 prevented its catalytic activity, raising the possibility that MYND-domain ligands could modulate EGLN1. However, we failed to detect an increased activity of EGLN1 on deletion of this region (Figure 3). Further studies will be required to establish the function of the EGLN1/EGLN2 unique N-terminal regions. Our results (Figure 3) also showed that a truncated version of EGLN1, lacking residues 76–177 from its N-terminal unique region, behaved as full-length EGLN1. We believe that this conclusion is important, since this construct has been extensively used to study EGLN biochemistry [6,10,11,14,16].

Another conclusion from our experiments is that the region sufficient for P402 binding (residues 201–252, shown in cyan and dark blue in Figures 5A and 5B) is relatively far from the catalytic site. The involvement of residues 201–252 in HIF α binding is supported by the presence of exposed hydrophobic patches in this region that were predicted to enable interaction of EGLN with its substrate [21]. Since substrate specificity for many enzymes is generally assumed to derive from features in close proximity to, or inside, the catalytic site, we find this observation significant. Analysis of the sequence determinants on HIF1 α ODDs revealed that residues remote to the target proline residue are required for optimal EGLN activity [11]. Thus, whereas the target proline residue localizes at the catalytic site, some of these ODD residues are distant enough to allow them to interact with EGLN regions relatively remote from the catalytic site.

Further analysis of the EGLN1 residues 201–252 led us to the identification of a minimal sequence (residues 236–252) necessary for NODD binding. This region contains the highest density of dissimilar residues between EGLN1 and EGLN3 (Figure 5E) and was previously suggested to have a role on substrate discrimination [21]. Since several of the differential residues are polar, the surface electrostatic potential of the finger-like region is very different for EGLN3 (mostly positive) and EGLN1 (containing positive and negative surface potential patches). Interestingly, sequence alignment reveals that, N-terminally to P402, there is an evolutionarily conserved cluster of basic residues (K389, K391 and K392) that is missing from the equivalent sequence upstream of P564. It is tempting to suggest that the ability of EGLN1 to bind P402 is mediated through the interaction of the negatively charged surface of its finger-like structure with this basic region. Conversely, since this negative surface potential seems to be absent from the equivalent region of EGLN3, as suggested by homology modelling, the interaction with P402 would be hindered in EGLN3. Hence, the combination of structural data and our present findings would support a model in which the presence of a mixed surface potential would allow interaction with both NODD and CODD in EGLN1, whereas the positive surface potential in the corresponding region of EGLN3 would favour CODD binding.

Given their role on HIF regulation, the EGLNs have been considered as potential drug targets. In fact, several inhibitors based on analogues of the reaction co-substrate 2-oxoglutarate have been described [27,28] and are widely used as research tools.

Figure 6 Effect of N-terminal region chimaeric constructs on HIF1α activity

(A) Schematic diagram representing the chimaeric constructs generated by exchange of the indicated regions between EGLN1 and EGLN3. Numbers indicate amino acid residue positions relative to the EGLN1 sequence. (B) HeLa cells were transfected with an HRE-driven reporter construct together with a plasmid encoding for HIF1α (none) or a combination of plasmids encoding for HIF1α and the indicated chimaeric enzymes. Results are normalized average reporter activities obtained in four independent experiments ($n = 4$). (C–E) Cells were treated as in (A), except that P402A (C), P564A (D) or P402A/P564A (E) mutant forms of HIF1α were used for the assays. The asterisks (***) indicate that mean values for N1/3A and U1N1C3 are significantly different ($P < 0.001$) from those obtained for U1N3C3. Differences between N1/3B and U1N3C3 were not statistically significant ($P > 0.05$). (F) Yeast cells were transformed with AD–HIF1α constructs together with DBD–EGLN chimaeric proteins. Serial dilutions of transformed clones were grown on plates lacking leucine and tryptophan (CONTROL) or plates lacking leucine, tryptophan, histidine and adenine (INTERACTION). The results shown are representative for two independent experiments. The abbreviations and symbols used are as described in for Figure 1.

However, the development of inhibitors based on analogues of 2-oxoglutarate that are specific for the individual EGLNs may be difficult to obtain, given their high degree of similarity within the catalytic site [21]. In this regard, our results open the possibility of devising specific inhibitors for the prolyl hydroxylases paralogues on the basis of their differential substrate-binding site identified herein.

A remaining open question is why EGLNs evolved different substrate-binding sites. It is tempting to speculate that different

EGLNs, particularly EGLN1 and EGLN2 as against EGLN3, might have a distinct substrate repertoire. In this regard, it is noteworthy that HIF1α, which is thought to be the evolutionary ancestor of HIFs gene family, has evolved two hydroxylation sites, whereas human HIF3α retains only CODD. On the other hand, EGLN1, which displays dual substrate specificity, seems to be the evolutionary origin of prolyl hydroxylases, whereas EGLN3 has evolved to recognize only CODD. It is therefore conceivable that this apparent selective pressure on certain EGLN HIF isoforms

could reflect specialized functions of EGLN paralogues in HIF regulation. Although this topic falls out of the scope of the preset study, an increased knowledge on the specifics of HIF regulation by the three EGLN paralogues, especially *in vivo*, should help in elucidating these appealing issues.

We thank Dr María Calzada (Servicio de Immunología, Hospital de la Princesa, Diego de León, Madrid, Spain) for critically reading this manuscript before its submission and for her valuable suggestions. This work was supported by grants from Ministerio de Ciencia y Tecnología (SAF2002-02344 and SAF2005-00180 to L. P. and SAF2004-0824 to M. O. L.), from RECAVA (Red Investigación Cardiovascular) and Comunidad Autónoma de Madrid (S-SAL-0311.2006).

REFERENCES

- Wenger, R. H., Stiehl, D. P. and Camenisch, G. (2005) Integration of oxygen signaling at the consensus HRE. *Sci. STKE* **306**, re12
- Schofield, C. J. and Ratcliffe, P. J. (2004) Oxygen sensing by HIF hydroxylases. *Nat. Rev. Mol. Cell Biol.* **5**, 343–354
- Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J. M., Lane, W. S. and Kaelin, Jr. W. G. (2001) HIF α targeted for VHL-mediated destruction by proline hydroxylation: implications for O₂ sensing. *Science* **292**, 464–468
- Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J. et al. (2001) Targeting of HIF- α to the von Hippel–Lindau ubiquitylation complex by O₂-regulated prolyl hydroxylation. *Science* **292**, 468–472
- Epstein, A. C., Gleadle, J. M., McNeill, L. A., Hewitson, K. S., O'Rourke, J., Mole, D. R., Mukherji, M., Metzen, E., Wilson, M. I., Dhanda, A. et al. (2001) *C. elegans* EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43–54
- Bruick, R. K. and McKnight, S. L. (2001) A conserved family of prolyl-4-hydroxylases that modify HIF. *Science* **294**, 1337–1340
- Ivan, M., Haberberger, T., Gervasi, D. C., Michelson, K. S., Gunzler, V., Kondo, K., Yang, H., Sorokina, I., Conaway, R. C., Conaway, J. W. and Kaelin, Jr. W. G. (2002) Biochemical purification and pharmacological inhibition of a mammalian prolyl hydroxylase acting on hypoxia-inducible factor. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13459–13464
- Taylor, M. S. (2001) Characterization and comparative analysis of the EGLN gene family. *Gene* **275**, 125–132
- Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., Wykoff, C. C., Pugh, C. W., Maher, E. R. and Ratcliffe, P. J. (1999) The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399**, 271–275
- Hirsilä, M., Koivunen, P., Günzler, V., Kivirikko, K. I. and Myllyharju, J. (2003) Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor HIF. *J. Biol. Chem.* **278**, 30772–30780
- Koivunen, P., Hirsilä, M., Kivirikko, K. I. and Myllyharju, J. (2006) The length of peptide substrates has a marked effect on hydroxylation by the HIF prolyl 4-hydroxylases. *J. Biol. Chem.* **281**, 28712–28720
- Masson, N., Willam, C., Maxwell, P. H., Pugh, C. W. and Ratcliffe, P. J. (2001) Independent function of two destruction domains in hypoxia-inducible factor- α chains activated by prolyl hydroxylation. *EMBO J.* **20**, 5197–5206
- Huang, J., Zhao, Q., Mooney, S. M. and Lee, F. S. (2002) Sequence determinants in hypoxia-inducible factor-1 α for hydroxylation by the prolyl hydroxylases PHD1, PHD2, and PHD3. *J. Biol. Chem.* **277**, 39792–39800
- Li, D., Hirsilä, M., Koivunen, P., Brenner, M. C., Xu, L., Yang, C., Kivirikko, K. I. and Myllyharju, J. (2004) Many amino acid substitutions in a HIF-1 α -like peptide cause only minor changes in its hydroxylation by the HIF prolyl 4-hydroxylases: substitution of 3,4-dehydroproline or azetidine-2-carboxylic acid for the proline leads to a high rate of uncoupled 2-oxoglutarate decarboxylation. *J. Biol. Chem.* **279**, 55051–55059
- Appelhoff, R. J., Tian, Y.-M., Raval, R. R., Turley, H., Harris, A. L., Pugh, C. W., Ratcliffe, P. J. and Gleadle, J. M. (2004) Differential function of the prolyl hydroxylases PHD1, PHD2 and PHD3 in the regulation of the hypoxia inducible factor. *J. Biol. Chem.* **279**, 38458–38465
- Landazuri, M. O., Vara-Vega, A., Viton, M., Cuevas, Y. and del Peso, L. (2006) Analysis of HIF-prolyl hydroxylases binding to substrates. *Biochem. Biophys. Res. Commun.* **351**, 313–320
- Gietz, D., St Jean, A., Woods, R. A. and Schiestl, R. H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acid Res.* **20**, 1425
- Aragones, J., Jones, D. R., Martin, S., San Juan, M. A., Alfranca, A., Vidal, F., Vara, A., Merida, I. and Landazuri, M. O. (2001) Evidence for the involvement of diacylglycerol kinase in the activation of hypoxia-inducible transcription factor 1 by low oxygen tension. *J. Biol. Chem.* **276**, 10548–10555
- Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cell by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752
- Chan, D. A., Sutphin, P. D., Yen, S.-E. and Giaccia, A. J. (2005) Coordinate regulation of the oxygen-dependent degradation domains of hypoxia inducible factor 1 α . *Mol. Cell. Biol.* **25**, 6415–6426
- McDonough, M. A., Li, V., Flashman, E., Chowdhury, R., Mohr, C., Lienard, B. M. R., Zondlo, J., Oldham, N. J., Clifton, I. J., Lewis, J. et al. (2006) Cellular oxygen sensing: crystal structure of hypoxia-inducible factor prolyl hydroxylase (PHD2). *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9814–9819
- Baker, D. and Sali, A. (2001) Protein structure prediction and structural genomics. *Science* **294**, 93–96
- Ozer, A., Wu, L. C. and Bruick, R. K. (2005) The candidate tumor suppressor ING4 represses activation of the hypoxia inducible factor (HIF). *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7481–7486
- To, K. K. W. and Huang, L. E. (2005) Suppression of hypoxia-inducible factor 1 α (HIF-1 α) transcriptional activity by the HIF prolyl hydroxylase EGLN1. *J. Biol. Chem.* **280**, 38102–38107
- Shen, C., Shao, Z. and Powell-Coffman, J. A. (2006) The *Caenorhabditis elegans rhy-1* gene inhibits HIF-1 hypoxia-inducible factor activity in a negative feedback loop that does not include *vhl-1*. *Genetics* **174**, 1205–1214
- Choi, K.-O., Lee, T., Lee, N., Kim, J.-H., Yang, E. G., Yoon, J. M., Kim, J. H., Lee, T. G. and Park, H. (2005) Inhibition of the catalytic activity of hypoxia-inducible factor-1 α prolyl-hydroxylase 2 by a MYND-type zinc finger. *Mol. Pharmacol.* **68**, 1803–1809
- Mole, D. R., Schlemminger, I., McNeill, L. A., Hewitson, K. S., Pugh, C. W., Ratcliffe, P. J. and Schofield, C. J. (2003) 2-Oxoglutarate analogue inhibitors of HIF prolyl hydroxylase. *Bioorg. Med. Chem. Lett.* **13**, 2677–2680
- Warnecke, C., Griethe, W., Weidemann, A., Jürgensen, J. S., Willam, C., Bachmann, S., Ivashchenko, Y., Wagner, I., Frei, U., Wiesener, M. and Eckardt, K.-U. (2003) Activation of the hypoxia-inducible factor pathway and stimulation of angiogenesis by application of prolyl hydroxylase inhibitors. *FASEB J.* **17**, 1186–1188
- Altschul, S. F., Gish, W., Miller, W., Meyers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410

Received 2 August 2007/23 August 2007; accepted 28 August 2007

Published as BJ Immediate Publication 28 August 2007, doi:10.1042/BJ20071052