# Attenuated function of a variant form of the helix-loop-helix protein, Id-3, generated by an alternative splicing mechanism

Richard W. Deed*, Michelle Jasiok, John D. Norton

*CRC Department of Gene Regulation, Paterson Institute for Cancer Research, Christie Hospital NHS Trust, Wilmslow Road, Manchester M20 9BX, UK*

Received 1 July 1996; revised version received 29 July 1996

**Abstract The Id family of helix-loop-helix proteins function as negative regulators of DNA binding, basic helix-loop-helix proteins in the regulation of cell growth and differentiation. We report here on the identification of a 17 kDa variant of the 14 kDa Id-3 protein termed Id-3L (long version) which possesses a unique 60 amino acid carboxy-terminus generated by read-through of a 'coding intron' and alternative splicing. Northern analysis revealed expression of a minor 1.1 kb Id-3L transcript together with the predominant 0.95 kb Id-3 transcript in the majority of adult human tissues analysed. The variant Id-3L protein is functionally distinguishable from conventional Id-3 since in in vitro DNA mobility shift assays, it was greatly impaired in its ability to abrogate binding of the basic helix-loop-helix protein, E47, to an E box recognition sequence.**

*Key words:* Gene expression; Transcription factor; Helix-loop-helix; RNA splicing

## 1. Introduction

Members of the Id family of helix-loop helix proteins are ubiquitously expressed and function at a general level as positive regulators of cell growth and negative regulators of differentiation [1]. Id functions are thought to be mediated primarily through the ability of this class of HLH protein to antagonise DNA binding and transcriptional activation by members of the somewhat larger family of basic HLH proteins, exemplified by the myogenesis-determining MyoD protein [2] and the E2A-encoded protein, E47, in B lymphopoiesis [3], through the formation of 'inactive' HLH heterodimers. In addition, Id gene expression is cell cycle regulated [4–6] and functional ablation experiments using antisense oligonucleotide blockade [6,7] and antibody microinjection [8] have suggested a role in G1 cell cycle progression, which may be intimately linked to control functions in cell growth/differentiation.

Whilst each of the four known Id genes (Id-1–4) has been shown to encode a single, 14–17 kDa protein [1], three laboratories have reported independently on a variant form of the prototype Id-1 protein (Id-IH [7], Id-1.25 [9], Id-1B [10] hereinafter referred to as 'Id-1L') generated by an alternative splicing mechanism which utilises an open reading frame in the first intron to generate proteins ranging in size from 13 to 17 kDa, with a distinct carboxy-terminus [9]. Rat Id-1L displays altered dimerisation properties in being able to form stable homodimers in contrast to the conventional 14 kDa Id-1 protein [9], and this has been suggested to represent a mechanism for attenuation of Id function, analogous to that described for the CREM [11] and *Drosophila* CF2 [12] transcription factor isoforms.

We previously reported that the human Id-3 gene displays a similar overall organisation of protein-coding sequences to that of Id-1, consistent with evolution from a common ancestral Id gene [13]. Specifically, the Id-3 gene harbors a 107 nucleotide intron interrupting the C-terminal protein coding sequence at an analogous position to the coding intron present in Id-1. Despite the disparity in primary sequence between the Id-1 and Id-3 genes in this region [13], these observations prompted an analysis of the potential of the Id-3 gene to encode a variant protein by a similar alternative splicing mechanism to that described for Id-1.

## 2. Materials and methods

### 2.1. Analysis of Id-3 mRNA

A commercial nylon membrane (Clonetech) loaded with 2 μg poly A mRNA from various human tissues was used for Northern hybridisation employing either a full length 0.95 kb human Id-3 cDNA probe [5] or a 107 nucleotide probe specific for the first intron of the human Id-3 gene. This latter probe was generated by PCR (25 cycles) of a cosmid clone using a 5′ primer, 96196 corresponding to nucleotides 1040–1061 in Deed et al. [5,13] and a 3′ primer, 96197 corresponding to nucleotides 1121–1145. The same 5′ primer was also used in combination with a 3′ RACE Tag primer (Clonetech) to generate a PCR fragment of approx. 0.65 kb from human placental cDNA. The identity of this and of the intron-specific PCR product was confirmed by nucleotide sequencing as described elsewhere [13]. $^{32}P$ labelling of probes, hybridisation and washing were performed essentially as described by Deed et al. [5]. Filters were re-probed with a glyceraldehyde phosphate dehydrogenase (GAPDH) probe to monitor equivalence of RNA loading.

### 2.2. Culture of human B cells and immunoprecipitation

Primary human B cells were isolated and purified from fresh tonsillar tissue essentially as described previously [14]. Briefly, after dispersal of tissue and lymphoprep separation (Flow Laboratories) T cells were depleted by incubation with neuraminidase-treated sheep red blood cells (Sigma) and further purified by lymphoprep centrifugation. Primary B cells and the cell line Reh (immature B cell) were cultured in RPMI medium supplemented with 10% FCS (Gibco). For [$^{35}S$]methionine labelling, cells were starved for 1 h in cysteine-methionine free medium (ICN-Flow Laboratories) containing 5% dialysed FCS and then incubated for an additional 2 h in fresh medium supplemented with 200 μCi/ml of 'Promix' (Amersham, spec. act. 1000 Ci mmol$^{-1}$). As appropriate, cells were stimulated with the phorbol ester, PMA (30 nM) for 5 h prior to harvest. Labelled cells were harvested by first being washed once in PBS and then lysed in RIPA buffer (50 mM Tris-HCl, pH. 8.0, 150 mM NaCl, 1% NP40, 0.1% SDS, 0.5% sodium deoxycholate), supplemented with the following protease inhibitors: 0.1 mM PMSF, 100 μg/ml each of leupeptin, bestatin, aprotinin and 2 mM levamisole.

Prior to immune precipitation, lysates were sheared by passage through a 19 g needle and particulate matter was collected by centrifugation at 2500 rpm for 3 min in a refrigerated microfuge. The cell lysate was pre-cleared by incubation with 10 μl of normal rabbit sera and 50 μl of protein A/G Sepharose beads (Santa Cruz Biotechnology,

*Corresponding author. Fax: (44) (0161) 446 3109.
E-mail: grgrwd@picr.cr.man.ac.uk


*PII* S0014-5793(96)00868-X

USA) for 30 min at 4°C with agitation. Supernatants were recovered by centrifugation and used in subsequent immune precipitations. For each immunoprecipitation reaction, 500 μl of cleared lysate was incubated with 5 μl of Id-3 specific antiserum, RD6 (raised against a full length Id-3 protein; Deed et al., unpublished) and rocked for 1 h at 4°C, after which 30 μl of protein A/G Sepharose beads was added and the incubation continued for a further 1 h. The immune complexed antigens were pelleted by centrifugation and washed four times with lysis buffer then finally re-suspended and boiled for 5 min in SDS sample buffer. The precipitated antigens were fractionated on a 12% SDS-polyacrylamide gel and fluorographed essentially as described previously [5].

*2.3. In vitro transcription and translation*

Constructs of Id-3 and Id-3L in the plasmid pcDNA3, (Invitrogen) were linearised with the enzyme *Xba*I and 2 μg was transcribed for 2 h at 37°C with T7 polymerase using a commercial kit (Promega UK) employing conditions recommended by the supplier. A 250 bp fragment of a truncated E47 cDNA [15] cloned into the pBluescript plasmid was linearised with the enzyme *Eco*RI for in vitro transcription with T3 polymerase. In vitro translation using a rabbit reticulocyte lysate system (Promega, UK) was carried out essentially according to the manufacturer's instructions typically using one tenth of the template RNA generated by in vitro transcription above. Radiolabelled protein was generated by performing parallel reactions incorporating [$^{35}$S]methionine (spec. act. 1000 Ci mmol$^{-1}$, Amersham, UK). Resulting proteins were analysed on 12.5% SDS-polyacrylamide gels or by DNA mobility band shift assay [5].

*2.4. DNA mobility shift assay*

A 20 bp double-stranded oligonucleotide containing the wild-type E box recognition sequence, ACCTGAA**CAGATGG**TCGGCT or mutant: ACCTGAA**CCGATT**GTCGGCT (the E-box element is shown in bold and differences are underlined), was used in all reactions to assay for E47 homodimer binding essentially as described previously [5]. Competition reactions with either wild-type or mutant probe (above) established the specificity of the reaction (data not shown). DNA mobility shift assays were carried out essentially as described by Murre et al. [16]. Oligonucleotides were self-annealed, 5′ end-labelled using [γ-$^{32}$P]ATP (spec. act. 3000 Ci mmol$^{-1}$, Amersham, UK) with polynucleotide kinase and subsequently purified on a polyacrylamide gel. Equimolar equivalence of Id-3 and Id-3L proteins was determined by [$^{35}$S]methionine incorporation and normalisation with respect to methionine content of the translated proteins. Approx. $4\times10^4$ cpm of labelled oligonucleotide was incubated in a final volume of 10 μl of programmed or unprogrammed rabbit reticulocyte lysate. Each binding reaction was carried out by first incubating the lysate at 37°C for 20 min in binding buffer comprising 10 mM Tris pH 7.4, 50 mM NaCl, 1 mM DTT, 1 mM EDTA. Subsequently, 1 μg of poly(dIC:dIC) (Sigma) was added as a non-specific competitor and incubated for a further 20 min. Finally, the probe was added and the incubation was continued for a further 20 min at 37°C prior to analysis on a 5% non-denaturing polyacrylamide gel essentially as described previously [5].

## 3. Results and discussion

Previously, we identified a 107 nucleotide intron located near the C-terminal coding region of the Id-3 gene in a region that is also interrupted by a short intron sequence in the Id-1 gene [13]. As shown in Fig. 1, the nucleotide sequence of this Id-3 intron predicts an open reading frame contiguous with that of the first 100 amino acids of the prototype Id-3 protein. The open reading frame continues beyond the intron sequence and into the second exon after which it terminates at a position just downstream of the prototype Id-3 termination codon. The size of the predicted protein, which we have designated 'Id-3L' (long version of Id-3) is 160 amino acids compared with 119 amino acids for the prototype Id-3 protein (Fig. 1) and, because the two predicted proteins utilise alternate open reading frames in the second exon region, they each have a unique C-terminus, 19 amino acids long in the case of Id-3 and 60 amino acids long for Id-3L.

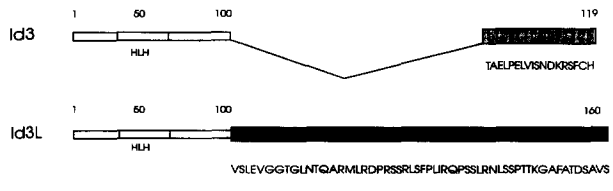


Fig. 1. Schematic representation of the structure and predicted amino acid sequences of Id-3 and the spliced variant Id-3L. Data for nucleotide/amino acid sequence were taken from [5,13]. The relative positions of the regions common to Id-3 and Id-3L, encompassing the HLH domain, are depicted by open boxes, and the amino acid sequences of the carboxy termini unique to each protein are shown.

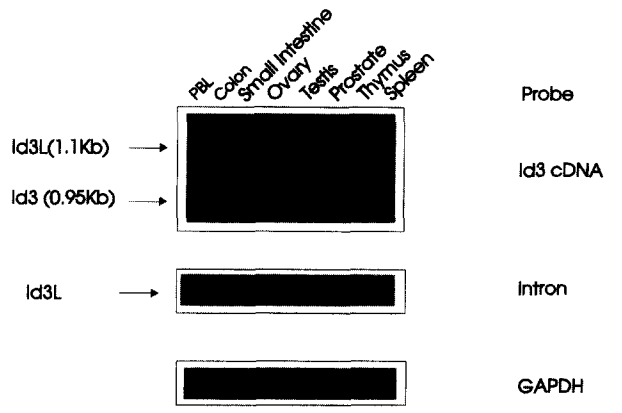


Fig. 2. Expression of Id-3 and Id-3L transcripts in human tissues. Replicate Northern blots were hybridised with either a full length Id-3 cDNA probe or with a 107 nucleotide, PCR-generated probe specific for the first intron of the human Id-3 gene. The membrane was re-probed with a GAPDH probe as a control for loading. Sizes of fragments were estimated relative to RNA markers (Clontech). PBL, peripheral blood lymphocytes (resting).

Northern blot analysis of RNA from primary tissues (Fig. 2) and from various cell lines (data not shown) using a full length Id-3 cDNA probe, revealed, in addition to the predominant 0.95 kb Id-3 transcript, a less abundant 1.1 kb transcript, consistent with an Id-3L mRNA generated by alternative splicing. As shown in Fig. 2, this 1.1 kb transcript (but not the 0.95 kb Id-3 transcript) was indeed detected with an intron-specific probe. PCR analysis of human placental cDNA using a primer flanking the 5′ boundary of the intron in combination with a 3′ RACE Tag primer followed by nucleotide sequencing (see Section 2) confirmed the presence of the alternatively spliced Id-3L transcript (data not shown). Most cell types/tissues expressing the 0.95 kb Id-3 transcript also expressed the 1.1 kb Id-3L mRNA, albeit at lower abundance. One exception was spleen, in which Id-3L was barely detectable (Fig. 2).

In common with most other early response genes, Id genes display a peak level of expression a few hours following mitogenic stimulation of quiescent cells, prior to the G1/S transition of the cell cycle [5–8]. Northern blot analysis for Id-3 and Id3-L encoding transcripts identified identical kinetics of induction following mitogenic stimulation of quiescent B cells (data not shown). To determine whether the Id-3L mRNA is translated in vivo, we examined Id-3 protein in stimulated primary human B cells which are particularly abundant in

this class of Id protein (Norton and Deed, unpublished observations). As shown in Fig. 3, resting ($G_0$) B cells barely yielded detectable Id-3 protein following [$^{35}$S]methionine labelling and immunoprecipitation with antibody raised against the full length Id-3 protein. Following mitogenic stimulation however, both the 15 kDa Id-3 protein together with a very minor 17 kDa species representing a candidate Id-3L protein, (poorly visible in Fig. 3) became detectable. Neither protein species was immune precipitable using pre-immune serum as a control (Fig. 3). In exponentially growing cells from the immature B cell line Reh, the candidate 17 kDa Id-3L protein was more easily detectable (Fig. 3).

The functional role of the alternatively spliced Id-1L protein is currently unknown [7,9,10]. The only reported difference between this and the prototype Id-1 protein, is in homodimerisation potential [9]. Whilst Id-1L readily forms stable homodimers, Id-1 does not and this has been suggested to represent a potential mechanism of negatively regulating Id function by sequestering Id-1L into an 'inert' homodimer configuration [9]. However, we recently found that the prototype Id-3 protein itself forms stable homodimers (in contrast to Id-1) by virtue of favourable charge-charge interaction at the dimer interface [17]. As an alternative functional property, we therefore compared the ability of Id-3 and Id-3L to form stable heterodimers with the bHLH protein E47, thereby abrogating binding to an 'E box' recognition sequence in band shift assays. Initially, in vitro translations were programmed to generate equivalent molar yields of the Id-3 and Id-3L proteins (Fig. 4A). Increasing inputs of these translates were then compared with an unprogrammed lysate (UP) for their ability to ablate the appearance of the E47 homodimer band shift as shown in Fig. 4B. In these experiments, E box binding to the USF protein which arises as an endogenous activity in reticulocyte in vitro translations (see [5]), provides a convenient internal control. Consistent with previous studies [4,5], the prototype Id-3 protein readily blocked E47 binding to DNA. However, translates programmed for Id-3L led to only a minor diminution of E47-E box binding at the highest input and even this was only marginally higher than that seen with unprogrammed lysate (Fig. 4B). Thus, we conclude that the ability of Id-3L to antagonise E47 binding to DNA is greatly impaired compared to that of the prototype Id-3.

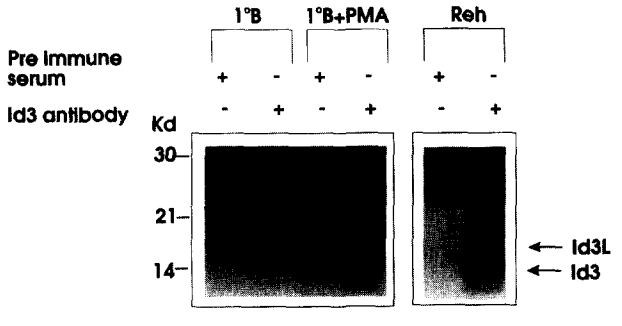


Fig. 3. Analysis of Id-3/Id-3L proteins in human B cells by immune precipitation. Purified human tonsillar B cells (1°B) were cultured in either the presence or absence of the phorbol ester PMA (30 nM) as indicated for 5 h. Cells were labelled with [$^{35}$S]methionine for 2 h then lysed and analysed by immune precipitation using either control pre-immune serum or anti-Id-3 (RD6) serum raised against a full length, bacterially synthesized Id-3 protein. Exponentially growing cells from the B cell line, Reh, were [$^{35}$S]methionine labelled for 2 h and similarly analysed by immune precipitation. Autoradiographs of 12% SDS-polyacrylamide gels are shown after exposure for 4 days.

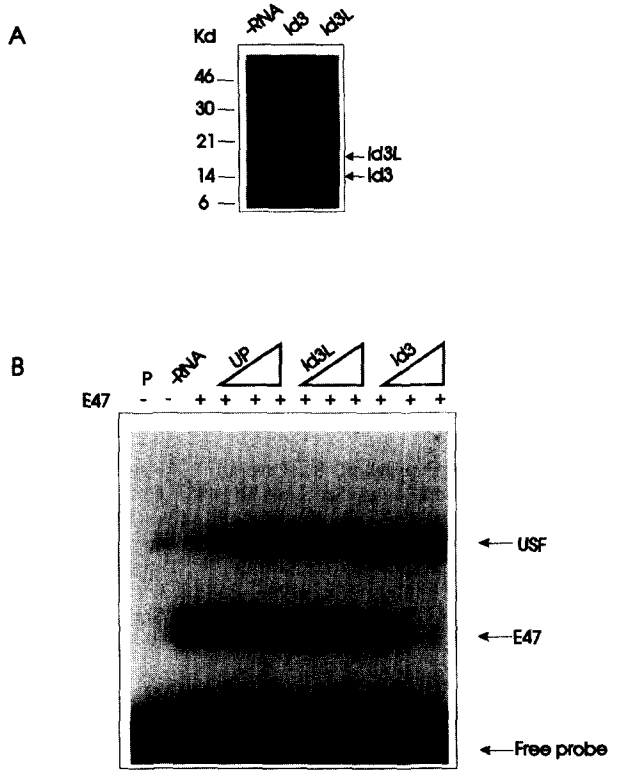


Fig. 4. Analysis of Id-3 and Id-3L function by DNA mobility shift assay. (A) In vitro translates of [$^{35}$S]methionine-labelled proteins either unprogrammed (−RNA) or programmed with Id-3 or Id-3L transcripts were analysed on a 12% polyacrylamide gel. (B) Band shift analysis using a $^{35}$P-labelled E box probe (control, lane P) after incubation with unprogrammed lysate (−RNA) or with in vitro translated E47 protein are shown. Incubations were challenged with increasing input (5, 10 and 15 μl) of competitor, either as control, unprogrammed lysate (UP) or with lysate programmed with Id-3 or Id-3L as indicated. In vitro translations of Id-proteins were performed in parallel with those shown in (A) except without labelled methionine and molar equivalence of Id-3 and Id-3L protein inputs were verified in further experiments (data not shown). The retarded complex labelled 'USF' arises from a contaminant in unprogrammed reticulocyte lysate (see control '−RNA' lane) and provides a convenient internal control [5].

Molecular modelling and site-directed mutagenesis studies have clearly established that the HLH domain is the primary determinant of dimerisation potential of Id (and bHLH) proteins [17–19]. Our data on Id-3L and that previously reported for Id-1L [9] now illustrate that domains well removed from the HLH region can have a dramatic effect on dimerisation function. Whilst the principal effect of generating an alternative carboxyterminus at least for the Id-3 protein appears to result in a 'loss of function' variant, as discussed previously for Id-1 [9,10], such alterations in dimerisation specificity may manifest in enhanced affinity for other bHLH targets which are not normally involved in dimerisation with conventional Id proteins. Alternatively, IdL proteins may target nuclear transcriptional effectors which do not belong to the bHLH family, such as the recently described MIdA1 binding partner for Id-1 [20], or the tumour suppressor protein, Rb which reportedly binds Id-2 [21].

*Acknowledgements:* We thank Dr. R. Watson for the gift of recombi-

nant E47 plasmid, Dr. J. Heighway for the multiple tissue Northern blot and Mrs. D. Johnston and Dr. G. Atherton for technical assistance. This work was supported by the UK Cancer Research Campaign and the Medical Research Council.

## References

[1] Littlewood, T.D. and Evan, G.I. (1995) Protein Profile, Transcription Factors 2: helix-loop-helix 2, 621–702.
[2] Benezra, R., Davies, R.L., Lockshon, D., Turner, D.L. and Weintraub, H. (1990) Cell 61, 49–49.
[3] Dorschkind, K. (1994) Cell 79, 751–753.
[4] Christy, B.A., Sanders, L.K., Lau, L.F., Copeland, N.G., Jenkins, N.A. and Nathans, D. (1991) Proc. Natl. Acad. Sci. USA 88, 1815–1819.
[5] Deed, R.W., Bianchi, S.M., Atherton, G.T., Johnston, D., Santibanez-Koref, M., Murphy, J.J. and Norton J.D. (1993) Oncogene 8, 599–607.
[6] Barone, M.V., Pepperkok, R., Peverali, F.A. and Philipson, L. (1994) Proc. Natl. Acad. Sci. USA 91, 4985–4988.
[7] Hara, E., Yamaguchi, T., Nojima, H., Ide, T., Campisi, J., Okayama, H. and Oda, K. (1994) J. Biol. Chem. 269, 2139–2145.
[8] Peverali, F.A., Ramqvist, T., Saffrich, R., Pepperkok, R., Barone, M.V. and Phillipson, L. (1994) EMBO J. 113, 3291–4301.
[9] Springhorn, J.P., Singh, K., Kelly, R.A. and Smith, T.W. (1994) J. Biol. Chem. 269, 5132–5136.
[10] Zhu, W., Dahmen, J., Bulfone, A., Rigolet, M., Hernandez, M.-C., Kuo, W.L., Puelles, J.L.R., Rubenstein, M. and Israel, A. (1995) Mol. Brain Res. 30, 312–326.
[11] Foulkes, N.S., Borrelli, E. and Sassone-Corsi, P (1991) Cell 64, 739–749.
[12] Hau, T., Gogos, J.A., Kirsh, S.A. and Kafatos, F.C. (1992) Science 257, 1946–1950.
[13] Deed, R.W., Hirose, T., Mitchell, E.L.D., Santibanez-Koref, M.F. and Norton, J.D. (1994) Gene 151, 309–314.
[14] Murphy, J.J. and Norton, J.D. (1993) Eur. J. Immunol. 23, 2876–2881.
[15] Voronova, A. and Baltimore, D (1990) Proc. Natl. Acad. Sci. USA 87, 4722–4726.
[16] Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Cabrera, C.V., Buskin, J.N., Hauschka, S.A., Lassar, A.B., Weintraub, H. and Baltimore, D. (1989) Cell 88 537–544.
[17] Wibley, J., Deed, R., Jasiok, M., Douglas, K. and Norton, J. (1996) Biochim. Biophys. Acta 1294, 138–146.
[18] Pesce, S. and Benezra, R. (1993) Mol. Cell Biol. 13, 7874–7880.
[19] Goldfarb, A.N., Lewandowska, K. and Shoam, M. (1996) J Biol. Chem. 271, 2683–2688
[20] Shoji, W., Inone, T., Yamamoto, T. and Obinata, M. (1995) J. Biol. Chem. 270, 24818–24825.
[21] Iaverone, A., Garg, P., Lasorella, A., Hsu, J. and Israel, M.A. (1994) Genes Devel. 8, 1270–1284.