Transcriptional regulation of cadherins during development and carcinogenesis

HÉCTOR PEINADO, FRANCISCO PORTILLO and AMPARO CANO*

Department of Biochemistry, Faculty of Medicine, Universidad Autónoma de Madrid, Instituto de Investigaciones Biomédicas “Alberto Sols”, CSIC-UAM, Madrid. Spain

ABSTRACT The cadherin superfamily of Ca²⁺-dependent homophilic adhesion molecules plays a critical role in regulating cell-to-cell interactions. During development, the expression of different cadherins is highly dynamic, since they are associated with the morphogenesis, establishment and/or maintenance of different tissues. Alterations in cadherin expression or function occur frequently during carcinogenesis, such as the loss of the epithelial cadherin (E-cadherin) and/or the aberrant expression of other cadherins. Indeed, the aberrant expression of cadherins has been detected during carcinoma invasion, a process which is reminiscent of the epithelial-mesenchymal transition (EMT) so important in many critical developmental processes. The functional regulation of cadherins can occur at many different levels, from transcriptional regulation to the control of the strength of the cadherin-mediated cell-cell interaction. In this review, we will focus on the transcriptional control of cadherin expression, both in development and carcinogenesis, paying particular attention to the regulation of E-cadherin given its proposed role as a suppressor of invasion. We will discuss the main genetic and epigenetic mechanisms involved in down-regulating E-cadherin expression, and we will analyse the mechanisms involved in regulating EMT, in an attempt to elucidate which elements are common to this process in both physiological and pathological situations.

KEY WORDS: cadherins, EMT, development, tumour invasion, transcriptional repression

Cadherins: a truly great family

The different cadherins constitute a superfamily of transmembrane receptors that mediate Ca²⁺-dependent cell-to-cell homophilic interactions. Since the description of the first member of this superfamily, uvomorulin (latter called E-cadherin), the cadherin superfamily has grown to include more than 80 members in both vertebrates and invertebrates (Nollet et al., 2000). The different members of this superfamily can be organised into several subfamilies according to their structural and/or functional organisation (Fig. 1). In this way, 5 different sub-families can be considered: 1) Classical cadherins (type I), mainly localised to adherens junctions; 2) Highly related type II cadherins; 3) Desmosomal cadherins (desmocollins and desmogleins) that form desmosomal junctions; 4) Protocadherins, mainly implicated in neural development; 5) Cadherin-related proteins, like the Flamingo and Fat-like cadherins (Nollet et al., 2000).

In general terms, the structure of the cadherins is highly conserved among the different members of the superfamily, particularly in terms of the extracellular domain that is formed by different numbers of the so-called cadherin domains (CD, or EC) (Takeichi, 1995). The classical cadherins present five cadherin domains (EC1 to EC5) which are responsible for the homophilic interaction of these molecules, especially the N-terminal EC1 domain. However, other sequences are also required for interactions between cadherins of different subfamilies. The association of calcium ions with a linker region connecting two of the EC domains induces the conformational changes necessary for the extracellular domain of cadherins to mediate their adhesive interactions. The cytoplasmic domain of cadherins is also conserved within each subfamily, and in the case of the classical cadherins, it is this domain that interacts with catenins which link the cadherin to the actin cytoskeleton.

Although the extracellular domain is sufficient to mediate cell-cell contacts, the cytoplasmic domain of cadherins is crucial to enhance the strength of this interaction and for cellular signaling.

An interesting discussion has arisen in recent years regarding

Abbreviations used in this paper: CID, CtBP interacting domain; CtBP, carboxy-terminal binding protein; EMT, epithelial-mesenchymal transition; HDAC, histone deacetylase; L-CAM, liver cell adhesion molecule; LMP1, latent membrane protein; MBPs, methyl binding domain proteins; MET, mesenchymal-epithelial transition; SNAG, snail-Gfi1 common domain.
the three dimensional distribution and interactions of cadherins. In the accepted model, based on the crystal structure of the EC1-EC2 domain in the classical E- and N-cadherins, cadherin dimers form dimers in the cytoplasmic membrane by the cis-clustering of monomers and lateral dimerisation. The cis-dimer expressed by adjacent cells can therefore interact in trans, adopting the characteristic cadherin-zipper structure that strengthens the adhesive junctions. However, more recent data obtained from the crystal structure of the five EC domains of C-cadherin has challenged the zipper model, suggesting an alternative in which the EC1 to EC3 domains of each cis-dimer are involved in the intercellular interactions (Boggon et al., 2002).

Cadherin expression during development

During embryonic development, each member of the cadherin superfamily exhibits its own specific spatio-temporal expression pattern. The elegant pioneering studies of Takeichi on E- and N-cadherin expression during mouse embryogenesis established a critical role for cadherins as regulators of morphogenesis (Takeichi, 1988; 1995). Both E- and N-cadherin are considered as prototypic cadherins in epithelial and mesenchymal tissues, respectively, and they are involved in the regulation of morphological events such as gastrulation, neurulation, cardiogenesis and somitogenesis. E-cadherin is the first adhesion molecule expressed in the mouse embryo, at the 8-cell stage, and is essential for the compaction of the morula and the subsequent organisation of epithelial tissues. Expression of E-cadherin is maintained in all epithelial tissues, but it is silenced during the process of EMT and in established mesenchymal cells (Takeichi, 1988; 1995). Loss of function studies (using functional antibodies, antisense nucleotides or transgenic knock-out mice for E-cadherin) have demonstrated that E-cadherin is crucial for early mouse development and the maintenance of epithelial morphology (Larue et al., 1996). Similar studies on N-cadherin have shown that it is required for cardiogenesis, somitogenesis and the correct development of the nervous system (Larue et al., 1996; Radice et al., 1997).

During early mouse development, dynamic and reciprocal changes in E- and N-cadherin expression can be detected (known as “cadherin switching”), for example when embryos undergo the process of EMT at the primitive streak where a switch from E- to N-cadherin expression occurs (Nakagawa and Takeichi, 1995). However, the E- to N-cadherin switch is not a generalised event when EMT occurs during development, and indeed it does not occur during the emigration of the neural crest cells from the neural tube in mouse embryos. Pre-migratory neural crest cells do not express either E- or N-cadherin, but N-cadherin can be detected in neural crest cells at their final destination (Takeichi, 1988; 1995). In addition, although they are embryonic lethal, N-cadherin null mutant mice are able to initiate gastrulation (Radice et al., 1997). It has been suggested that other cadherins (such as cadherin-11 (OB-cadherin) and/or cadherin-6B) could be involved in the regulation of cell-cell interactions in the absence of N-cadherin, raising the hypothesis of partial functional substitution between some cadherin members (Radice et al., 1997).

The expression of cadherins during EMT in other tissues has also been examined, such as during the formation of the chick neural crest. In these cells, L-CAM (liver-cell adhesion molecule, the avian homologue of E-cadherin) is substituted by N-cadherin and cadherin-6B during neural plate invagination (Nakagawa and Takeichi, 1995). Furthermore, during migration of the chick neural crest, no N-cadherin or cadherin-6B can be detected in these cells, but rather they express cadherin-7 which is maintained until they reach their appropriate destinations (Nakagawa and Takeichi, 1995). Similarly, cadherin switching has been observed in the organisation of other embryonic tissues, supporting the suggestion that N-cadherin is implicated in cell differentiation, segregation and migration (Radice et al., 1997).

Cadherins are also involved in the developmental process that is the inverse of EMT, Mesenchymal–Epithelial Transition (MET), which occurs during events such as in kidney morphogenesis. Indeed, cadherin-6 (K-cadherin) is the main adhesion molecule involved in the conversion of mesenchymal cells into an epithelium in the kidney. Studies with cadherin-6 deletion mutants have confirmed that R-cadherin can compensate for the lack of cadherin-6 during MET in the kidney, but in R-cadherin deficient mice this process is barely affected (Dahl et al., 2002). Double knock-out mice for R-cadherin and either P-cadherin, N-cadherin or cadherin-6 show no substantial changes in phenotype when compared to the single null mutants, suggesting that the regulation of cell-cell contacts during MET in kidney cells is complex and involves the participation of multiple cadherin molecules (Dahl et al., 2002). Additionally, the atypical Ksp-cadherin (cadherin 18) is also involved in epithelial kidney morphogenesis, this cadherin contains a highly truncated cytoplasmic domain and is specifically expressed in renal tubular epithelial cells.

This overview of cadherin expression during developmental processes indicates that the up- or down-regulation of specific cadherin members is usually highly controlled and is associated with the gain or loss of cell-cell adhesion. This suggests that control of cadherin expression and function during EMT or MET might involve at least two levels of regulation: 1) Control of cadherin expression at the transcriptional level; and 2) Control of cadherin function through signalling pathways that influence their stabilisation.

Cadherin expression during tumour progression

Modifications of cell-cell interactions similar to those observed during development also occur in some pathological situations, such as during carcinogenesis, where the disruption of cell-cell contacts is one of the key events in tumour progression. During the last decade, the expression of cadherins during tumour progression has been the subject of extensive studies, the majority of which have focused on the role of E-cadherin. The loss of E-cadherin expression and/or function has been observed during tumour progression of most carcinomas (reviewed in Takeichi, 1993; Birchmeier and Behrens, 1994), and this has been related to the induction of EMT which frequently occurs during carcinoma invasion (Christofori and Semb, 1999; Thiery, 2002). These observations, together with data gathered from other model systems, indicate that as initially proposed, E-cadherin can be considered as an invasion suppressor gene (Birchmeier and Behrens, 1994). Other cadherins have also been implicated in tumour progression; for example, it has been proposed that cadherin-13 (T-cadherin, H-cadherin) acts as a tumour suppressor gene since its expression is significantly diminished in several types of carcinomas. Silencing of cadherin-13 seems to occur by a combination of allelic loss and promoter hypermethylation (Toyokawa et al., 2001). However, since cadherin-13 lacks the intercellular
domain necessary for binding to cytoplasmic proteins and potential signalling pathways, the molecular mechanisms by which it might act as a tumour suppressor are still unclear.

Recent evidence has indicated that N-cadherin overexpression can be correlated with an increase in the invasive potential of breast cancer cells and that this induction appears to be correlated with the decrease in the expression of other cadherins, such as E- and P-cadherin in squamous cell carcinoma lines. This situation is reminiscent of the E- to N-cadherin switching observed in some developmental processes. Moreover, N-cadherin seems to have a dominant effect, enhancing the motility of tumour cells even in presence of E-cadherin in breast cancer cells (Nieman et al., 1999). Neo-expression of N-cadherin has also been observed in other tumour types, such as prostate (Tomita et al., 2000) and intestinal gastric carcinomas (Rosivatz et al., 2002), suggesting a functional role for this cadherin in cancer progression.

The increase in invasion and metastasis associated with N-cadherin expression could be explained by a destabilisation of cell-cell adhesion mediated by N-cadherin that allows the cells to detach from the primary tumour and associate selectively with the stroma and/or endothelium. Increased secretion of MMP-9 and activation of the MAPK pathway after N-cadherin over-expression, or the functional co-operation between N-cadherin and FGFR signalling observed in breast cancer cells (Suyama et al., 2002), could also account for the association between N-cadherin and invasiveness. Nevertheless, while down-regulation of E-cadherin expression is often correlated with EMT in many carcinoma cells, over-expression of N-cadherin is not. Taken together, these observations suggest that aberrant expression of N-cadherin might play an active role in tumour progression only in some specific carcinoma cells or tumours.

Cadherin-11 is normally expressed in the brain, yet it has also been proposed to be upregulated during tumour progression. Several studies have reported over-expression of cadherin-11 in breast and prostate cancer cell lines, and in some types of gastric and prostate tumours including prostate metastatic lesions (Tomita et al., 2000). Moreover, an increase in cadherin-11 expression is frequently associated with the up-regulation of N-cadherin, suggesting that rather than the loss of E-cadherin, the switching on of both these cadherins could be important for the acquisition and/or maintenance of an invasive and metastatic phenotype in specific cancer cell types (Tomita et al., 2000).

Fig. 1. Organisation of the cadherin superfamily. The main structural and functional characteristics of the cadherin subfamilies and types are indicated. Representative examples are shown, with gene denomination of some of them in brackets. The main structural domains and schematic structure of the different cadherin types are shown at the right.
The aberrant expression of other members of the cadherin family has also been observed in carcinomas, as is the case for cadherin-6 whose over-expression is correlated with malignancy in renal cell carcinomas, independent of E-cadherin loss. Indeed, it has been proposed that this molecule could be used as an independent marker for the prognosis of tubular cell carcinomas. Recently, N-cadherin and cadherin-8 have also been implicated in the development and invasiveness of renal cell carcinomas. In contrast, several studies on P-cadherin have shown that a decrease in the expression of this cadherin can be correlated with the invasiveness of lung carcinoma and melanomas. However, a recent study also revealed that a switch between E- and P-cadherin expression might be involved in the progression from Stage I to Stage II ovarian cancer (Patel et al., 2003). Moreover, aberrant P-cadherin expression has been detected in several carcinomas, including breast, cervical, oesophagus and basal cell carcinomas.

In general, the cadherin switching that has been observed in many tumours seems to be a mechanism to “make new friends”, provoking a shift to a more dynamic adhesion state through the expression of new cadherins. This change in the adhesion state may allow tumour cells to migrate, as well as promoting interactions with the stroma and endothelia. However, the mechanisms underlying cadherin switching in development and cancer currently remain unclear. The heterogeneous pattern of switching between several cadherin members observed in different types of tumours and carcinoma cells suggest that environmental cues and/or specific signalling pathways might underlie cadherin switching. All this information indicates that by elucidating the mechanism that regulate the expression of distinct cadherin members in

---

**Fig. 2. Schematic representation of the E-cadherin promoter region and E-box sequences in mouse, human and dog.**

(A) The E-cadherin promoter region is represented with its representative proximal control elements in the proximal region (exerting either a positive or negative effect on E-cadherin expression). The CAAT, GC and E-box 1/E-box 3 are conserved in all three promoters, while the E-box 2 is only present in the mouse promoter region. An additional E-box (E-box 4) downstream of the transcription initiation site is present in the human promoter, but is not conserved in the region encompassing the mouse promoter. Binding of identified factors to the different E-boxes is shown in the lower part.

(B) Sequence of the E-boxes present in the proximal region of the mouse, human and dog E-cadherin promoter. Observe the high degree of conservation of E-box 1 and E-box 3.
different cancer types, we might obtain important clues to further understand tumourigenesis.

**Mechanisms of E-cadherin regulation during tumour progression**

Most of our current knowledge regarding the way in which cadherin expression is controlled during tumour progression has derived from studies on E-cadherin. Indeed, the involvement of E-cadherin in epithelial morphogenesis and homeostasis, and its proven anti-invasive role in carcinoma progression has fomented interest in understanding the regulatory mechanisms that control E-cadherin expression under normal and pathological circumstances.

The human E-cadherin gene (CDH1) is located at the 16q22.1 locus and is comprised of 16 exons, spanning 99 Kb of genomic DNA. The mouse E-cadherin gene (cdh1) has a similar exon-intron organisation and is located on chromosome 8. Loss of E-cadherin function is thought to be one of the most important events during carcinoma progression (Birchmeier and Behrens, 1994), an assumption that is reinforced by the fact that germ line mutations in CDH1 induce a predisposition to diffuse-type gastric cancer (see Strathdee, 2002).

During tumour progression, E-cadherin can be functionally inactivated by different mechanisms, including somatic mutation and down-regulation of gene expression through promoter methylation and/or transcriptional repression. Indeed, down-regulation of the CDH1 gene expression is observed in many carcinomas (Birchmeier and Behrens, 1994), while inactivating mutations, frequent in diffuse gastric carcinomas and in lobular breast carcinomas, are rarely observed in other types of tumours (reviewed in Berx et al., 1998; Strathdee, 2002). Indeed, the most frequent genetic alterations of CDH1 in tumours involve exon-skipping and out-of-frame mutations, and in most cases mutations occur in combination with a loss of heterozygosity of the wild type allele (Berx et al., 1998)

**Epigenetic mechanisms of E-cadherin silencing**

Anomalous hypermethylation of promoter CpG islands is an important means of repressing tumour suppressor genes. DNA methylation is known to be involved in the development of human cancers, often characterised by a generalised hypomethylation of DNA and local hypermethylation of CpG islands in the promoters and upstream exons of many genes. The hypermethylation of CpG islands is associated with the recruitment of methyl DNA binding proteins (MBDs) and of histone deacetylase activity (HDACs), which together contribute to the compaction of the DNA in the promoter region and hence, to gene inactivation.

The initial studies on CDH1 promoter methylation not only established a relationship between E-cadherin silencing and the methylation of CpG islands in several carcinoma cell lines, but also observed that treatment with 5-azacytidine re-activated E-cadherin expression. Furthermore, methylation of the E-cadherin promoter was observed in primary prostate and breast tumour lesions, in contrast to normal adjacent tissues where it remained unmethylated. Since these two pioneering studies, a large amount of data regarding E-cadherin promoter methylation and tumour progression has been accumulated from a huge number of cell lines and tumours (reviewed in Strathdee, 2002). Of particular interest was the finding that in diffuse-type gastric cancer, methylation of the E-cadherin promoter serves as a second hit to somatic or germ-line genetic alterations in inactivating E-cadherin expression (Grady et al., 2000). Indeed, it has also been proposed that methylation of E-cadherin CpG islands can act as a second hit for transcriptional repression to silence E-cadherin in ductal breast carcinoma (Cheng et al., 2001).

The origin of abnormal E-cadherin promoter methylation is still obscure. In the case of gastric cancer, exogenous agents like oxygen species and H. pylori infection have recently been invoked as the basis for ageing-related methylation in the stomach. In support of this hypothesis, E-cadherin methylation has been considered as an early event in gastric carcinogenesis initiated by H. pylori infection, and a progressive increase in E-cadherin methylation has been correlated with ageing. However, a relationship between H. pylori infection and E-cadherin methylation during ageing is still not clear (Tamura, 2002). Nevertheless, a series of studies have suggested a relationship between viral oncogenesis and aberrant methylation, as seen in Epstein Barr Virus (EBV)-positive gastric carcinomas where the E-cadherin promoter, among others, is hypermethylated (Kang et al., 2002).

Despite the loss of E-cadherin observed in primary carcinomas, E-cadherin is frequently re-expressed at secondary metastatic foci and even within lymph node metastasis. This indicates that E-cadherin expression is dynamically regulated during tumour progression, a fact that apparently contradicts the assumed irreversibility of DNA methylation. Analysis of E-cadherin methylation in breast carcinomas and in other tumour types has shown that aberrant hypermethylation of CpG islands in the promoter region of E-cadherin often occurs prior to invasion, indicating that it is an early event in tumourigenesis. Moreover, hypermethylation of the E-cadherin promoter seems to persist in invasive and metastatic lesions in breast carcinomas, but with a rather heterogeneous pattern between tumour cells. Additionally, dynamic changes in E-cadherin methylation have been proposed to occur during tumour progression (Graff et al., 2000). According to this model, tumour cells with predominantly methylated E-cadherin alleles would be responsible for the initiation of metastatic behaviour, but a decrease in methylation, leading to the re-expression of E-cadherin, might occur in the established metastatic lesions and could contribute to survival of the cells in breast cancer. Indeed, re-expression of E-cadherin in cells in the blood stream and secondary tissues was earlier suggested as a general means for cancer cells to survive (Takeichi, 1993). Whether dynamic E-cadherin methylation also occurs in other tumour types, thereby explaining changes in E-cadherin expression during tumour progression, remains to be established. On the other hand, methylation of the E-cadherin promoter is not always correlated with E-cadherin silencing (Strathdee, 2002), indicating that additional genetic or epigenetic modifications might be required for E-cadherin down-regulation. The molecular mechanisms involved in regulating E-cadherin methylation are starting to be unravelled. Recently, the association of several MBDs, such as MeCP2, MBD1 and MBD2, with the E-cadherin promoter have been reported in several cancer cell lines (Koizume et al., 2002). Interestingly, methylation of E-cadherin CpG islands and MeCP2 expression have been shown to cooperatively and epigenetically regulate E-cadherin expression in colorectal cancer cells (Darwanto et al., 2003). It is also possible...
that other epigenetic mechanisms, like histone deacetylation, participate in regulating E-cadherin expression. Histone H3 deacetylation is associated with CpG methylation-mediated gene silencing in other contexts and recently, the participation of both epigenetic modifications was reported for E-cadherin down-regulation in several cancer cell lines (Koizume et al., 2002).

**Transcriptional regulation of E-cadherin silencing**

Besides regulation of E-cadherin by promoter hypermethylation and/or genetic alterations, direct transcriptional control of E-cadherin has emerged in the last years as an important regulatory mechanism of E-cadherin expression. The mouse E-cadherin promoter was first isolated in 1991. Its initial characterisation showed it was a TATA-less promoter containing several potential proximal regulatory elements, including a CCAAT box (-65), a GC-rich region (-30 to -58) and a palindromic element (-70 to -90) composed of two adjacent E-boxes flanked by four inverted nucleotides called ‘E-pal’ (Behrens et al., 1991). The proximal CCAAT and GC-rich regions are required for basal E-cadherin expression and are recognised by constitutive AP2 and Sp1 transcription factors, and CAAT-binding proteins, respectively (Behrens et al., 1991; Hennig et al., 1996; Faraldo et al., 1997). The E-pal element was initially described as an epithelial-specific regulator (Behrens et al., 1991), but subsequent studies showed it to be an active repressor in E-cadherin deficient cells (Hennig et al., 1996; Faraldo et al., 1997; Rodrigo et al., 1999). In addition, an epithelial-specific enhancer does exist in the first intron of the mouse E-cadherin gene, which is recognised by AP2 factors and other potential regulators (Hennig et al., 1996). An additional regulatory element, containing Ets-binding sites, has also been described in the mouse promoter and may contribute to the repression of E-cadherin expression (Rodrigo et al., 1999).

A comparison of the human, mouse and dog E-cadherin promoters (Comijn et al., 2001) showed that the CCAAT-box and GC-rich regions are conserved at similar locations (see Fig. 2). The proximal E-box in the mouse E-pal element is also conserved in the human and dog promoters, as is the additional E-box (E-box 3) located at -30 in all promoters. However, the additional E-box (E-box 4) downstream of the transcription initiation site in the human promoter (Batlle et al., 2000) is not present in the mouse promoter.

Several studies have demonstrated that the E-boxes in the proximal E-cadherin promoter repress its expression. Point mutations abolishing the two E-boxes in the mouse E-pal, or E-box 1 and 3 in the human promoter, produced a strong induction of E-cadherin promoter activity in different carcinoma cell lines deficient in E-cadherin expression. Evidence has also been presented that the E-box 4 also represses E-cadherin expression (Batlle et al., 2000; Hajra et al., 2002), hence, factors able to bind to the E-boxes and repress E-cadherin expression must be present in these cells. In vivo analysis of the endogenous human and mouse E-cadherin promoter by footprinting supports this hypothesis. The E-pal or E-box 1 were specifically protected in E-cadherin deficient cells, as well as in several dedifferentiated carcinoma cell lines and fibroblasts (Hennig et al., 1996; Rodrigo et al., 1999). An important concept derived from such studies was that the repressors bound to proximal E-boxes were able to overcome the positive effects of constitutive factors interacting with the basal regulatory elements of the E-cadherin promoter, such as AP2 and Sp1 (Rodrigo et al., 1999).

**Characterisation of E-cadherin repressors**

A major breakthrough in understanding the regulation of E-cadherin transcription has been the identification of several E-cadherin repressors. Using an approach based on the yeast one-hybrid system, we searched a cDNA library from E-cadherin deficient NIH3T3 fibroblasts for repressors binding to the E-boxes in the E-pal element of the mouse E-cadherin promoter. The specificity of the factors identified was confirmed in a parallel screen with an E-pal element in which a point mutation abolished the repressor function (Hennig et al., 1996; Faraldo et al., 1997) and which should also preclude the binding of regulatory factors. More than 90% of the identified factors that specifically bound to the E-boxes corresponded to transcription factors, of which, two in particular were highly represented: 49% corresponded to the zinc finger factor Snail and 32% to the class I bHLH factors, E47 (Cano et al., 2000; Pérez-Moreno et al., 2001). The remaining factors corresponded mainly to an additional class I bHLH factor, mITF2 (also called E2-2; F. Portillo, L. Holt and A. Cano, unpublished), while a single clone corresponded to another member of the Snail family,Slug (Bolós et al., 2003).

Functional characterisation of Snail indicated that it does indeed act as a strong repressor of the mouse and human E-cadherin promoters (Batlle et al., 2000; Cano et al., 2000). Snail repressor activity is strictly dependent on the intact E-pal element in the mouse promoter (Cano et al., 2000), but it apparently requires the three E-boxes of the human promoter (Batlle et al., 2000). Importantly, overexpression of Snail in epithelial cells produces a dramatic EMT, and the acquisition of migratory and in vitro invasive behaviour (Cano et al., 2000). Consistent with this behaviour, Snail and E-cadherin expression is inversely correlated in different mouse and human carcinoma and melanoma cell lines (Batlle et al., 2000; Cano et al., 2000; Poser et al., 2001). Initial studies of Snail expression in primary mouse skin carcinomas showed that it is specifically expressed in invasive areas, coinciding with the down-regulation of E-cadherin expression (Cano et al., 2000), thus strongly supporting the involvement of Snail in E-cadherin repression at the invasive front. Recent studies in breast carcinoma biopsies have highlighted the association between Snail expression, dedifferentiation and lymph-node status of ductal breast carcinomas (Cheng et al., 2001; Blanco et al., 2002). Similar observations have also been made recently in invasive hepatocellular carcinomas (Sugimachi et al., 2003). In breast carcinomas, Snail expression could be observed in a subset of ductal tumours without positive lymph nodes, suggesting its potential utility as an early prognostic marker of malignancy (Blanco et al., 2002).
human promoters. While the factors that act on the mouse promoter seem to preferentially interact with the E-pal element (including Snail, Slug and E47: Cano et al., 2000; Pérez-Moreno et al., 2001; Bolós et al., 2003), in the human promoter either all three E-boxes or E-box 1 and 3 appear to be required for the interaction of Snail/Slug or δEF1/SIP1, respectively (Batlle et al., 2000; Grootecaes and Frisch, 2000; Comijn et al., 2001; Hajra et al., 2002). These observations suggest that at least two of the three proximal E-boxes in the mammalian E-cadherin promoter are functionally similar in terms of recruiting repressors, regardless of their relative location. Interestingly, all of the E-cadherinrepressors identified exhibit expression patterns in developing embryos compatible with this function (see below).

Co-operation of the different repressors or a functional hierarchy in E-cadherin down-regulation?

One important question that remains unresolved when considering the transcriptional repression of E-cadherin relates to the specific participation of the different repressors in the down-regulation of E-cadherin during tumour progression. Analysing the expression of different repressors in several carcinoma cell lines has rendered some apparently contradictory results. Thus, in a panel of breast carcinoma cell lines, the inverse correlation between Slug expression and E-cadherin down-regulation was stronger than with Snail expression, suggesting that in vivo Slug may be a more likely repressor of E-cadherin breast cancer (Hajra et al., 2002). In contrast, when primary ductal breast carcinomas were analysed, Snail appeared to be the candidate repressor of E-cadherin expression (Cheng et al., 2001; Blanco et al., 2002). Similarly, Snail was associated with the down-regulation of E-cadherin in invasive hepatocarcinomas (Sugimachi et al, 2003), while other studies showed that E-cadherin down-regulation was associated with the co-expression of Snail and SIP1, or the preferential expression of SIP1 in different carcinoma cell lines (Comijn et al., 2001). Interestingly, an inverse correlation between E-cadherin and δEF1 (associated with Snail expression) but not with SIP1 has also been observed (Guaita et al., 2002). Furthermore, in several carcinoma and melanoma cell lines the inverse expression of E12/E47 and E-cadherin has been reported (Pérez-Moreno et al., 2001).

Some insights into the role of the different E-cadherin repressors have been obtained from in vitro studies. A comparative analysis of the binding affinities for the E-pal element of three E-cadherin repressors, Snail, Slug and E47, revealed that Snail binds with a higher affinity than the other two repressors. The affinity constant of Snail binding to the E-pal (0.5 nM) is one and two orders of magnitude higher than that of E47 and Slug, respectively (Bolós et al., 2003). This indicates that a hierarchy might exist in terms of the participation of the three repressors when present in the same biological context: Snail will predominate over E47, which in turn will predominate over Slug in silencing E-cadherin. An analysis of the expression of the different repressors in tumour biopsies should clarify the specific participation of each factor in down-regulating E-cadherin. In this regard, a recent study in gastric carcinomas has shown an association between E-cadherin repression and Snail expression in diffuse gastric carcinomas while SIP1expression was linked to E-cadherin

Fig. 3. A model for the regulation of E-cadherin expression involving different repressors and epigenetic mechanisms. A schematic illustration of the putative epigenetic mechanisms involved in E-cadherin regulation is shown. In this model, active transcription is probably achieved through the interaction of several activators, and involves the acetylation of histone residues which promotes the chromatin to adopt an open conformation. Under the influence of the appropriate signals, several repressors and corepressor molecules are recruited (depending on either the tissue or cellular context), attracting the repressor machinery which includes corepressor molecules, HDACs and methylases, and still unidentified molecules which will modify the acetylation and methylation status of chromatin, promoting E-cadherin silencing (see text for details).
repression in intestinal type gastric carcinomas (Rosivatz et al., 2002). Although still preliminary, these data indicate that the different repressors may participate in silencing E-cadherin specifically in different types of tumour or, perhaps, at defined stages of tumour progression. A more detailed analysis of the expression of these repressors in large-scale tumour samples is clearly needed to study this hypothesis, an aim that is presently hindered by the lack of good specific antibodies to the different repressors.

**Mechanism of E-cadherin repression: potential co-operation with epigenetic modifications**

The information available regarding the different mechanisms involved in E-cadherin silencing make it difficult to define a simple model where E-cadherin expression is regulated by just a single genetic, epigenetic or transcriptional control mechanism. It seems more likely that a combination of different mechanisms is responsible for defining the status of E-cadherin expression during tumour progression. As such, a modification of the classical “two hit hypothesis” for E-cadherin silencing has recently been proposed by Cheng et al., in which a collaboration between genetic, epigenetic and transcriptional control of E-cadherin expression is at play in ductal breast cancer (Cheng et al., 2001). In this system, Snail expression is correlated with the silencing of E-cadherin and hypermethylation of its promoter, rather than by methylation itself. Indeed, one interpretation of this model might invoke an interesting role for Snail in the co-ordination of both processes (Cheng et al., 2001).

The specific mechanisms by which Snail, and/or other repressors, mediates E-cadherin silencing are still largely unknown. The modification of chromatin by the co-ordinated action of DNA and/or histone methylation, and acetylation, has emerged as one of the major mechanisms for regulating the transcriptional activity of different regulatory genes. Indeed, several years ago the modification of chromatin was proposed as a mechanism to regulate E-cadherin expression, and the participation of histone deacetylase activities (HDACs) has been used to explain the repressor activity of human Slug on artificial promoters (Hemavathy et al., 2000).

Very recently, details of the mechanism underlying Snail repression of E-cadherin were revealed, providing a link between transcriptional control and the epigenetic modifications of the E-cadherin promoter (Peinado et al., 2004). Snail repression of the E-cadherin promoter involves the recruitment of a repressor complex formed, at least, by the co-repressor mSin3A, HDAC1 and HDAC2 (Peinado et al., 2004) (Fig. 3). The recruitment of this complex is mediated by the N-terminal SNAG domain of Snail, previously thought to act as the repressor domain. The presence of this complex results in a net decrease in the amount of acetylated histones H3/H4 and an increase in methylated K9 of histone H3 in the endogenous E-cadherin promoter. In turn, this leads to a compact organisation of the chromatin (Peinado et al., 2004). A preliminary analysis in the mouse skin carcinogenesis model indicates that Snail expression is also associated with the hypermethylation and silencing of the E-cadherin promoter, further supporting the notion of a connection between Snail and epigenetic modifications (Fraga, M., Peinado, H. et al., in preparation; see Fig. 3). The molecular mechanisms involved in this link remain to be elucidated, but the implication of MBDs and HDACs is an interesting aspect of E-cadherin silencing for further study, since treatment with inhibitors of DNA methylation and HDACs promotes the re-expression of E-cadherin (Peinado et al., 2004, and unpublished results).

The molecular mechanisms responsible for Slug repression remain to be established. Nevertheless, it is worth bearing in mind that apart from the N-terminal SNAG domain, Slug contains a partially conserved domain that interacts with the CtBP co-repressor (Hemavathy et al., 2000), opening the possibility for Slug to co-operate with different repressor complexes. CtBP interacting domains (CID) are present in the homeodomain-zinc binding factors δEF1 and SIP1, and an interaction with CID was previously reported to be necessary for δEF1 to effectively repress the E-cadherin promoter (Grootelaes and Frisch, 2000). The recent

---

**Fig. 4. Model of the participation of different E-cadherin repressors during tumour invasion.** In this model, several repressors like Snail and SIP1 would play a role in the first step of tumour progression leading to the initiation of the invasive process through the down-regulation of E-cadherin and other epithelial markers, and the induction of mesenchymal markers which promote invasion. Other repressors, like Slug or E-47, may contribute to the maintenance of the malignant phenotype by sustained E-cadherin repression and promoting ECM degradation and cell motility, provoking the appearance of isolated cells capable of eventually undergoing metastasis (see text for details).
isolated CtBP co-repressor complex containing different methylase and HDAC activities, as well as the δEF1 and SIP1 factors, and which is functionally active on E-cadherin promoter regulation (Shi et al., 2003), strongly support the participation of CtBP complexes in down-regulating E-cadherin expression (Fig. 3). Whether δEF1 and SIP1 factors do indeed play a direct or indirect role in the CtBP repressor complex remains to be established.

The mechanisms involved in E47-mediated repression of E-cadherin are still unknown. In contrast to Snail and Slug, preliminary analyses do not suggest that HDAC activity is involved but rather point to a regulatory mechanism involving heterodimerisation with either the class II bHLH factors, such as Twist, or Id regulators (E. Cubillo, H. Peinado, and A. Cano, unpublished).

In summary, current evidence indicates that silencing of E-cadherin transcription requires the participation of several repressor factors that interact with specific E-boxes in the proximal promoter. The repression of gene expression either involves the local modification of chromatin organisation through the recruitment of specific co-repressor complexes (Snail, Slug and potentially δEF1/SIP1), or heterodimerisation of repressors with specific partners (E47). These mechanisms of repression add an additional level of complexity to our understanding of the different factors that specifically repress E-cadherin expression. Thus, as well as the specific expression patterns and/or binding affinities for similar regulatory elements, repressor function might finally be regulated by the presence of, and/or affinity for, different co-repressors in specific cells or tissues. Thus, the regulation of expression and function of the different repressors needs also to be taken into account to obtain a comprehensive view of the transcriptional control of E-cadherin. In this sense, a recent study has elegantly shown that silencing of Snail expression can be triggered by the ligated oestrogen receptor in breast cancer cell lines, involving the specific participation of the MTA3 protein in an additional repressor complex (Fujita et al., 2003). In this system, Snail expression is derepressed by a deficiency in oestrogen receptors or in the MTA3 protein, producing the suppression of E-cadherin expression (Fujita et al., 2003).

Interestingly, epigenetic modifications of the Snail promoter also seem to be involved in the regulation of its expression during mouse skin tumour progression (Fraga, M., Peinado, H. et al., in preparation). This indicates that several co-repressor molecules and epigenetic mechanisms may also regulate the expression of epithelial and mesenchymal genes. Furthermore, Snail expression is induced by activated ILK (Tan et al., 2001) and by TGFβ, and/or oncogenic H-ras signals in epithelial cells (Peinado et al., 2003), and the nuclear localisation of Snail also seems to be under fine control (Dominguez et al., 2003). This information points to other important mechanisms that might influence the expression of E-cadherin via external signals during tumour progression.

Mechanisms underlying the regulation of other cadherin genes, or cadherin switching during tumour progression are poorly understood. A few additional cadherin promoters have been characterised, like P-, N- and Ksp-cadherin. All of them are TATA-less promoters and contain conserved CCAAT and GC-rich regions, as well as several E-boxes, with the exception of P-cadherin (Faraldo et al., 1997; Li et al., 1997; Whyte et al., 1999). Although some regulatory elements that confer tissue specific expression have been described in these promoters, potential elements responsible for activation or repression during tumourigenesis remain to be discovered. However, it is tempting to speculate that some of the E-cadherin repressors identified could also work as activators of other cadherin genes, as well as of mesenchymal genes.

**Regulatory mechanisms of E-cadherin expression in development. What can we learn about the role of repressors?**

As discussed above, E-cadherin is expressed in a highly dynamic fashion during early mouse development. It is present in all epithelial cells from the early morula stage but disappears in cells undergoing EMT or in established mesodermal cells (Takeichi, 1988; 1995). The mechanisms regulating E-cadherin expression during development are still poorly understood. Intriguing, a recent study was unable to find proximal upstream or downstream regulatory elements that might be responsible for the tissue specificity of E-cadherin expression, suggesting a major role for downstream tissue-specific enhancer elements (Stemmler et al., 2003). Similar studies on other cadherins suggest that the tissue specific transcriptional regulation of widely expressed cadherin members, like N-cadherin, is highly complex in vivo, probably involving different enhancer-specific regulators (Li et al., 1997). In contrast, the tissue specific regulation of cadherins whose expression is more restricted, like Ksp-cadherin in tubular kidney cells, is provided by proximal elements (Whyte et al., 1999).

Despite the lack of information on cadherin regulation in vivo, we now have a more comprehensive knowledge of the role of E-cadherin repressors in development. Snail family members are now recognised as important regulators of EMT during early vertebrate development (Nieto, 2002). Indeed, in mouse development the expression pattern of Snail is closely correlated with the down-regulation of E-cadherin expression in regions where EMT occurs, while Slug expression is observed in the already migratory mesenchymal cells (Cano et al., 2000; Nieto, 2002). The pattern of E2A/Cδ expression (encoding E12/E47 factors), and of E2-2δ (encoding mITF2), is similar to that of Slug (Pérez-Moreno et al., 2001; Holt et al., unpublished). On the other hand, Zfhx1b (encoding SIP1) is expressed in the neuroepithelium and both the premigratory and migratory neural crest cells of specific regions, as well as in the branchial arch mesoderm (van de Putte et al., 2003). Therefore, the developmental expression patterns of known repressors of E-cadherin are fully consistent with their repressor role in epithelial cell systems.

Analysis of Snail knockout embryos has provided compelling evidence for its role as a repressor of E-cadherin in EMT. Snail null mice present early embryonic lethality, and they fail to complete gastrulation, giving rise to a deficient mesodermal layer in which E-cadherin expression is maintained. In contrast, Slug null mice are viable, undergoing an apparently normal program of mesoderm and neural crest cell development (see Nieto, 2002). The absence of mesoderm and neural crest defects have also been observed in null mice for the E2A or E2-2A gene, and these phenotypes have been explained by the potentially redundant role of the class IbHLH factors encoded by both genes (E12/E47 and E2-2A/E2-2B, respectively). On the other hand, the Zfhx1b knockout mice are embryonic lethal at 9.5 dpc, and they exhibit several defects in neural crest migration or specification. Specifically, these mice are deficient in post-mitotic vagal neural crest, display defects in
cranial neural crest delamination and exhibit ectopic E-cadherin expression in the neuroepithelium (van de Putte et al., 2003).

The distinct expression patterns observed between the different E-cadherin repressors during development and their null phenotypes, suggest a hierarchy of these factors. Thus we can envisage a prominent role for Snail in inducing EMT, a more restricted role for SIP1 in migration and specification of distinct neural crest cell populations, and a role for Slug and E12/E47 (and E2-2 products) in the maintenance of the mesoderm. These developmental analyses strongly support the idea that the regulation of E-cadherin expression is similar during normal development and tumour progression (Cano et al., 2000), and support the specific and hierarchical participation of the different factors in the repression of E-cadherin during invasion.

Based on the available information we propose a model for the participation of the different E-cadherin repressors in tumour invasion (Fig. 4). Thus, the transient expression of Snail, or SIP1, will be involved in inducing the invasion process, with perhaps the more specific participation of each repressor in distinct types of carcinomas. In contrast, other repressors, such as Slug, E47 and perhaps E6F1, will be involved in the maintenance of the migratory invasive phenotype. Curiously, this model also fits with the relative binding affinities of some of the repressors (Snail, Slug and E47) as determined in vitro for the regulatory E-boxes, and could also explain some of the apparent discrepancies found in the literature. Indeed, it may become apparent that the breast carcinoma cell lines used in different analyses represent distinct stages of tumour progression and/or invasiveness and, therefore, the expression patterns of the different repressors observed, such as Snail and Slug or SIP1, could reflect their distinct origins. In other words, we need to think of tumour invasion as a highly dynamic process, similar to EMT in development, whereas we have probably so far only been looking at still images.

Although the information gathered so far has been instrumental to increase our understanding of how E-cadherin expression and tumour progression are regulated, a closer view of similar developmental processes, together with a large scale analysis of tumour biopsies, will certainly help to increase our understanding of the tumour invasion process. This will in turn contribute to the design of new therapeutic approaches based on the inhibition of the expression and/or function of E-cadherin repressors.

Acknowledgements

The authors apologize to those investigators whose experimental work has not been cited or cited indirectly due to space limitations. The authors thank the members of A. Cano’s laboratory for their excellent work and useful discussions. Experimental work at A. Cano’s laboratory was supported by grants from the Spanish Ministry of Science and Technology (SAF2001-2819), the Instituto de Salud Carlos III (FIS01/1174 and FIS03/03-10) and by Justesa Imagen, S.A.

References


