Journal of Cellular Biochemistry 61:531–542 (1996)

Structural and Functional Diversity of Cadherin Superfamily: Are New Members of Cadherin Superfamily Involved in Signal Transduction Pathway?

Shintaro T. Suzuki

Departments of Ophthalmology and Microbiology, University of Southern California School of Medicine and Doheny Eye Institute, Los Angeles, California, 90033

Abstract A large number of cadherins and cadherin-related proteins are expressed in different tissues of a variety of multicellular organisms. These proteins share one property: their extracellular domains consist of multiple repeats of a cadherin-specific motif. A recent structure study has shown that the cadherin repeats roughly corresponding to the folding unit of the extracellular domains. The members of the cadherin superfamily are roughly classified into two groups, classical type cadherins proteins and protocadherin type according to their structural properties. These proteins appear to be derived from a common ancestor that might have cadherin repeats similar to those of the current protocadherins, and to have common functional properties. Among various cadherins, E-cadherin was the first to be identified as a Ca^{2+}-dependent homophilic adhesion protein. Recent knockout mice experiments have proven its biological role, but there are still several puzzling unsolved properties of the cell adhesion activity. Other members of cadherin superfamily show divergent properties and many lack some of the expected properties of cell adhesion protein. Since recent studies of various adhesion proteins reveal that they are involved in different signal transduction pathways, the idea that the new members of cadherin superfamily may participate in more general cell-cell interaction processes including signal transduction is an intriguing hypothesis. The cadherin superfamily is structurally divergent and possibly functionally divergent as well.

Key words: cadherin superfamily, signal transduction pathway, adhesion proteins, evolution, biological role, structure

INTRODUCTION

In multicellular organisms, each cell is connected to other cells and to the extracellular matrix, physically as well as functionally, in a highly ordered manner to form and maintain the integrity of the organisms. It is quite obvious that some molecules must mediate these connections, and many investigators have been working in this research area for the past couple of decades. As a result, various so-called adhesion molecules have been isolated and characterized. The studies have revealed that most of these adhesion molecules belong to one of the following four protein families: the immunoglobulin family, the integrin family, the cadherin family, or the selectin family [for review, see Hynes and Lander, 1992].

Since the pioneering work by Holtfreter and associates [1955], the cell-cell adhesion mechanism has attracted the interest of many investigators, and cadherins were subsequently identified as a group of Ca^{2+}-dependent cell-cell adhesion proteins. Among these proteins, E-cadherin or uvomorulin was identified first as a molecule involved in the compunction process of the blastomere. Since then, although two other cadherins, N-cadherin and P-cadherin, were discovered soon after the identification of E-cadherin, the results of the extensive study of E-cadherin provide a paradigm of cadherin research [for review, see Takeichi, 1991; Geiger and Ayalon, 1992].

These three cadherins were the only members of cadherin family for some time, despite circumstantial evidence suggesting the existence of others. For the last 5 years or so, however, the whole picture has changed and various other cadherins and cadherin-related proteins have been identified [Koch et al., 1990; Mahoney et al., 1991; Ranscht and Dours-Zimmermann, 1991; Suzuki et al., 1991; Sano et al., 1993; Dantzig et al., 1994]. This protein family is now

Received March 18, 1995; accepted December 7, 1995.

Address reprint requests to Shintaro T. Suzuki, Doheny Eye Institute, 1450 San Pablo Street, DVRC-309, Los Angeles, CA 90033.

growing rapidly and has turned out to be a superfamily.

In this paper, I will describe some recent progress in the field of cadherin research. The emphasis will be on the new members of the cadherin superfamily. I will not necessarily describe the established concepts in this field; rather I will express my personal views. These views may differ from the established views in many issues, but I believe this attempt may provoke thoughts about the function of various members of the cadherin superfamily.

Sequence Features of Cadherin Superfamily

Recent cloning studies have revealed the basic sequence features of the cadherin superfamily and have provided the basis for the classification of the family members into four groups: classical cadherins, desmosomal cadherins, protocadherins, and other cadherin-related proteins. These proteins all share one property: their extracellular domains consist of multiple repeats of a cadherin-specific motif (cadherin repeat). This motif is approximately 110 amino acids long and contains several highly conserved short amino acid sequences and well conserved amino acids as shown in Figure 1. The number of cadherin repeats varies from four to more than thirty among different members of the cadherin superfamily. A closer look at the cadherin repeats shows that the repeat features are different from one group to another, and even from one repeat to another within one protein. Thus, each repeat of each family member has its own characteristic properties.

It can be difficult to determine the start and end sites of the cadherin repeats, because of their intrinsic cyclic nature. In this paper I will use the old definition of the repeating unit, described initially by Hatta et al. [1988] and refined further by our group and others [Mahoney et al., 1991; Sano et al., 1993; Tanihara et al., 1994a].

In contrast to the extracellular domains, the cytoplasmic domains of cadherin superfamily members are highly variable and contain different sequences.

(a) Classical cadherin family. The overall structure of classical cadherins is essentially the same. The signal sequence at the N-terminal site is flanked by a prosequence that contains a protease processing signal sequence, $^{K}/_{R}$RXKR, at the C-terminal end. The proteolytic cleavage at this site appears to be necessary for the activation of the classical cadherins. Following the prosequence, the extracellular domain is located at the N-terminal side and the cytoplasmic domain is at the C-terminal side. Both domains are connected with a single transmembrane segment. The cadherin extracellular domain consists of five repeats of a cadherin motif, each repeat showing characteristic features in addition to the common properties. The first cadherin repeat from the N-terminus (EC1) appears to contain the cell adhesion site [Blaschuk et al., 1990; Nose et al., 1990]. Cadherin EC3 contains $D^{Y}/_{F}E$ sequence instead of DRE sequence in the middle of the repeats and a characteristic one amino acid deletion near the end of the repeat. Another distinctive feature of classical cadherins is found in the EC5s. Classical cadherin EC5s have four characteristic cysteine residues, and DRE and DXNDNXPXF sequences are missing from the repeats. The cytoplasmic domains of classical cadherins show highly conserved sequences of about 150 amino acids. The C-terminal region is especially well conserved, suggesting the functional importance of this region, whereas the N-terminal half is relatively variable except for one small region near the N-terminus.

The alignment of the amino acid sequences of various cadherins shows that classical cadherins can be classified into two types [Tanihara et al., 1994a]. Type I classical cadherins (type I cadherins) include E-cadherin, M-cadherin, N-cadherin, P-cadherin, and R-cadherin or cadherin-4. The other type, Type II classical cadherins (type II cadherins), include the relatively newly found cadherins, such as cadherin-5 through cadherin-12. Many characteristic amino acids and short amino acid deletions and additions are present only in one of these two groups. Furthermore, the amino acid identity values among classical cadherins of the same group are higher than those between the two groups.

T-cadherin or cadherin-13 is a unique member among the cadherin superfamily in that it lacks the cytoplasmic domain and a part of the transmembrane domain [Ranscht and Dours-Zimmermann, 1991; Tanihara et al., 1994a]. However, this protein is linked to the cell membrane via a phosphoglycolipid. Despite this characteristic structure, the extracellular domain contains mostly the features of classical cadherin extracellular domains, especially type I cadherins. T-cadherin is likely to be a special form of the classical cadherin. Another interest-

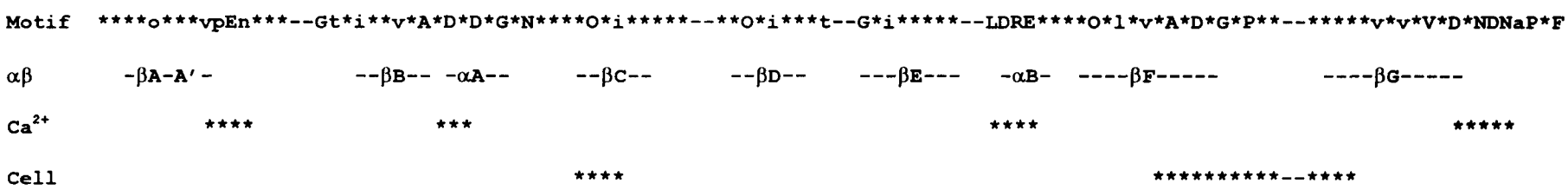

Fig. 1. Cadherin motif and possible cell-binding and Ca^{2+}-binding sites. Cadherin motif is described in **Motif** line. The *dashes* in the motif line indicate the regions of cadherin repeats showing variable lengths and *o* denotes aromatic amino acids. **Ca^{2+}**, **Cell**, and **αβ** lines describe the possible Ca^{2+}-binding site, homophilic binding site, and secondary structure elements, respectively [Oberduin et al., 1995].

ing form was found in cadherin-8, a type II cadherin. Cadherin-8 is expressed in two major forms in tissues: one form has the typical structure of classical cadherins; the other form has a truncated structure [Kido et al., unpublished observation]. In contrast to the case of T-cadherin, the truncated form of cadherin-8 ends near the N-terminus of the EC5 and entirely lacks the transmembrane domain and the cytoplasmic domain. It is unknown whether this form of protein is also produced for other members of the cadherin superfamily.

(b) Desmosomal cadherin family. The overall structure of desmosomal cadherins is very similar to that of classical cadherins [for review, see Amagai, 1995]. The major difference between the two is in the cytoplasmic domains. The cytoplasmic domains of desmosomal cadherins are longer than those of classical cadherins and have weak homology with those of classical cadherins. The unique cytoplasmic sequences appear to cause these proteins to interact with the intermediate filaments, either directly or indirectly, at desmosomes. Interestingly, desmoglein 1, a desmosomal cadherin, lacks the EC5 subdomain, and has only four cadherin repeats in its extracellular domain [Koch et al., 1990]. However, the EC3 has the characteristic features found in the EC3s of classical cadherins. Desmocollins have alternatively spliced forms that are not common in classical cadherins. The splicing occurs in the cytoplasmic domains, which may affect their interaction properties with the cytoplasmic proteins.

(c) Protocadherin family and protocadherin-related proteins. This family was discovered relatively recently in mammals, but the total number of family members appears to be very large [Sano et al., 1993]. No one has extensively searched for protocadherins or protocadherin-related proteins in various invertebrates yet, but protocadherin-related proteins are already known in various invertebrates such as *Drosophila,* nematodes, sea urchins, hydra, and planaria [Mahoney et al., 1991; Sano et al., 1993; Miller and McClay, personal communication; Pettitt et al., personal communication]. In contrast to the vertebrate protocadherins, the protocadherin-related proteins of invertebrates have not been studied extensively. However, it is likely that every multicellular organism expresses protocadherins or protocadherin-related proteins.

The overall structure of vertebrate protocadherins and invertebrate protocadherin-related proteins is similar to that of classical cadherins, but they share characteristic features that are not found in classical cadherins. These proteins do not have prosequences in contrast to classical cadherins. The extracellular domains contain more than five cadherin repeats that are very similar to each other in length and sequence properties. These repeats are similar to EC2s or EC4s of classical cadherins, but none of the repeats has the characteristic features of classical cadherin EC3s and EC5s. In mammals, the entire coding sequences for four different protocadherins have been determined [Sano et al., 1993; Sago et al., unpublished observation]. The alignment of these sequences suggests that each repeat of protocadherins has characteristic features as do the classical cadherins.

Another interesting feature of this family is its highly divergent cytoplasmic domain sequences, which suggest that this family is a heterogeneous one containing various subfamilies. Indeed, one subfamily of protocadherin-3 and its related protocadherins has already been identified in mammals [Sago et al., unpublished observation]. Also of interest is the finding that some protocadherin-related proteins from invertebrates contain sequences showing significant homology with those of classical cadherins [Oda et al., 1994; Miller and McClay, personal communication]. By analogy to the case of immunoglobulin superfamily (see below), some members may even contain sequences showing homology with those of protein kinase or protein phosphatase, although no such finding has yet been reported.

One recent study revealed a series of unique protocadherins that share the same sequence in their cytoplasmic domains, whereas the other part of the proteins are homologous to, but distinct from each other [Obata et al., unpublished observation]. How the corresponding mRNAs are formed is unknown. They may be produced by an alternative splicing or by some other mechanism. Whatever the mechanism is, this type of mRNA is unique and has not been reported before. It is possible that the resultant proteins have different specificity of interaction at the extracellular domains but have the same interaction with the cytoplasmic protein at the cytoplasmic domains. Therefore, these proteins could play an interesting role. At present, we do not know whether other protocadherins and/or

any other proteins have similar molecular forms. Thus, the protocadherin family appears to be a large one and to be structurally divergent.

(d) Other members of the cadherin superfamily. Two unique cadherin-related proteins (HPT/LI cadherins) were recently found in liver, intestine or kidney [Dantzig et al., 1994]. The extracellular domain of these proteins contains clear cadherin repeats. The distinctive property of these proteins is that while the basic cadherin repeat number is five, this protein contains two additional repeats. If we disregard the repeats, the extracellular domain structure is essentially the same as that of classical cadherins. However, the cytoplasmic sequences are very short and there is no homology with other members of the cadherin superfamily.

Several research groups have reported that an oncogene called *ret* contained cadherin repeats. However, the sequence similarity is very low, and the sequence lacks the many highly conserved amino acids and short amino acid sequences found in the cadherin repeats. It may not be the cadherin homologue or it may be a far-related cousin of the cadherin superfamily.

Since this field is still young, it is possible that more cadherin-related proteins will be found in various tissues or cells in the near future.

Evolution of Cadherin Superfamily

One fundamental question regarding the cadherin superfamily is whether the cadherin repeats are derived from one primordial cadherin motif or if they are simply the result of a convergent evolution. Each cadherin repeat is different from one protein to another; however, the constitutive elements of the cadherin motif are highly conserved in all the repeats of the cadherin superfamily members (cadherins), and the lengths between the well conserved elements are almost identical among various repeats of different cadherins of a variety of organisms from *C. elegans* to human. These findings favor the common origin hypothesis.

As described above, cadherins can be roughly classified into two groups by the cadherin repeat properties. One group includes classical cadherins, desmosomal cadherins, and HPT/LI cadherins. These proteins contain cadherin repeats that show the characteristic features of the classical cadherin EC3s and EC5s. The other group contains protocadherins and protocadherin-related proteins. The extracellular domains consist of highly homologous cadherin repeats and do not contain the characteristic EC3s and EC5s found in classical cadherins. Classical cadherins have not been identified outside the vertebrates so far, despite various efforts. Instead, two research groups have reported protocadherin-related proteins with cytoplasmic sequences that are homologous to those of classical cadherins [Oda et al., 1994; Miller and McClay, personal communication]. This finding provides the evidence for the possible evolutional connection between the two groups. It is likely that various primordial protocadherins had different cytoplasmic domains and played different roles. One group with classical cadherin-like cytoplasmic domains evolved to the current classical cadherins, desmosomal cadherins and HPT/LI cadherins were diversified during this process. The other groups became the current protocadherins (Fig. 2). If the above hypothesis is correct, the cadherin repeats found in current protocadherins appear to retain many properties of the primordial cadherin motif; hence, the name protocadherins.

However, one controversial result has been reported. The exon-intron structure for each cadherin repeat of classical cadherins is not the same [Miyatani et al., 1992]. We think that this difference in gene structure is due to the secondary modification of the genes during the evolution of classical cadherins. The genomic sequences for the extracellular domains of protocadherins and protocadherin-related proteins do not contain many introns; at least the genes examined so far contain no intron over the range of several cadherin repeats, which supports our hypothesis [Mahoney et al., 1991; Pettitt et al., personal communication; Kitagawa et al., unpublished observation]. Furthermore, the alternatively spliced form of cadherin-8 suggests that an intron is present at the site corresponding to the alternative splicing site near the N-terminus of EC5 that is not present in type I cadherin genes, although we have not determined the genomic structure of cadherin-8.

The comparative study of various cadherins may provide a clue for the study on structure-function properties of cadherins.

Higher Structure of Extracellular Domains of Cadherin Superfamily

Since the fundamental motif is essentially the same among various cadherins, the cadherin repeats appear to be important units both structurally and functionally. A structure model of

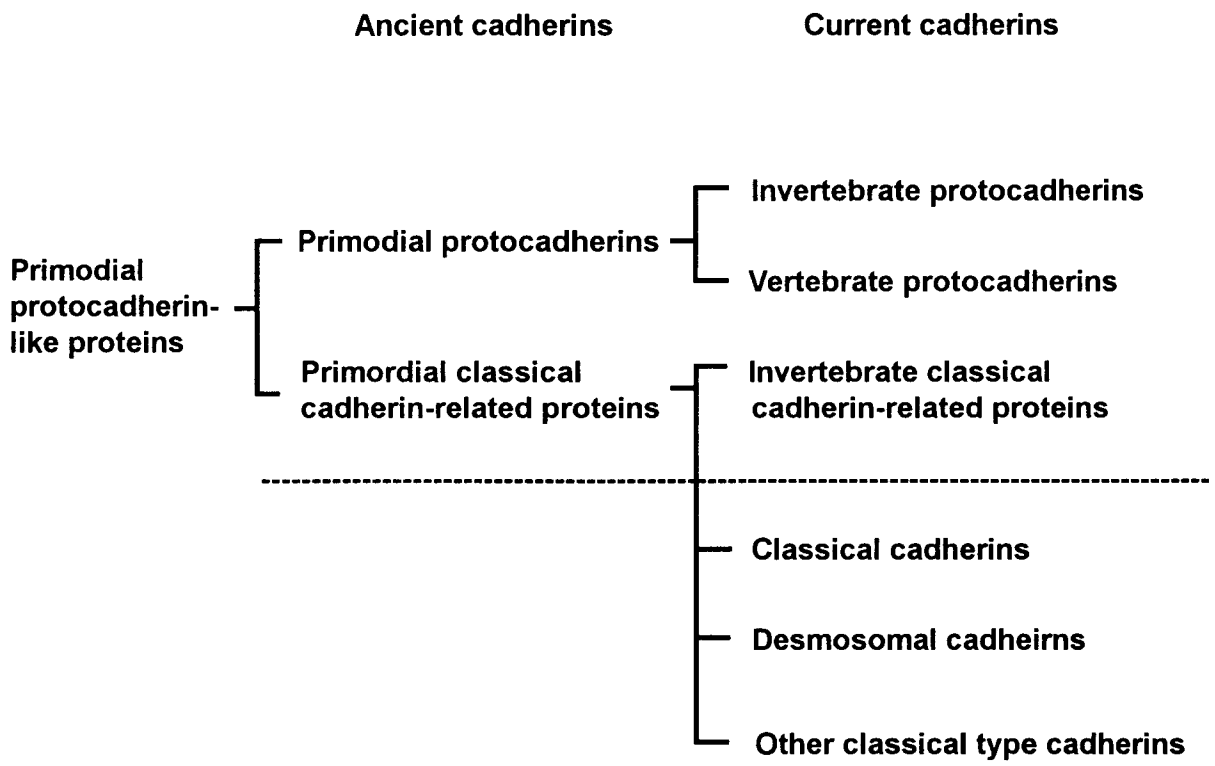

Fig. 2. Possible evolution of cadherin superfamily. Classical type cadherins are shown below the broken line and the other types of cadherins are shown above the broken line.

the cadherin repeats is essential for the study of the cadherin superfamily. A long-awaited structure model of cadherin repeats, obtained from the EC1 subdomain of mouse E-cadherin, has recently been reported [Overduin et al., 1995]. The message of this model is that the cadherin repeats indeed roughly correspond to the folding unit that forms an immunoglobulin-like structure. The lengths between the well conserved amino acids and short sequences are highly conserved among various cadherin repeats, and consequently the repeats consist of well conserved sequence blocks. Importantly, these sequence blocks basically correspond to the secondary structure components of the model (see Fig. 1). The postulated homophilic binding site is located on one side of the folded repeat and a Ca^{2+} binding site is formed between the two adjacent repeats.

This may be the turning point of structure-function study of the cadherin superfamily. It should also be noted, however, that the model may have several limitations. Since the EC1 sequence has several unique features among various cadherin repeats, the structure model would need some minor modification before it could be applied to other repeats. More importantly, this model does not tell much about the configuration of multiple repeats, especially in the acting state. Nevertheless, the overall structure is likely to be the same as the model among various cadherin repeats of different cadherins. The model should answer many structural and functional questions. Indeed, this model is consistent with various results obtained thus far and should prove useful for the experimental design of further studies.

At present, the evolutional relationship between the cadherin motif and the immunoglobulin motif is unclear. However, the two protein families appear to have many analogous properties.

(a) Homophilic binding site. Several sequences in EC1 have been suggested for the homophilic binding site and/or the specificity-determining site of cell adhesion [Blaschuk et al., 1990; Nose et al., 1990]. The structure model shows that these sites are close to each other at a side area of the folded structure; thus the antibodies against these regions inhibit the cell adhesion activity. Of course, this does not necessarily mean that all these sequences are directly involved in the cell adhesion activity and/or the specificity determination. If this model of the homophilic binding site is correct, however, then we can conclude that the binding site is not as

simple as the RGD sequence found in the integrin system. This is a reasonable conclusion, since the RGD sequence is just a ligand sequence that binds to a receptor, whereas the binding sites of cadherins correspond to receptor sites themselves. Clearly, the binding site and/or the specificity-determining site appears to be present in this narrow region. Ozawa et al. [1990], reported an antibody that recognized the EC5 sequence of E-cadherin inhibited the cell adhesion activity. One explanation might be that EC1 and EC5 are close enough in the homophilically bound structure that the antibody inhibits the adhesion activity, possibly by steric hindrance. Unfortunately, we do not currently have any concrete information regarding the quaternary structure of the cadherin extracellular domain. Certainly, this will be the next big challenge for the study of the structure of cadherins.

Blaschuk et al. [1990], reported that the HAV sequence near the C-terminal region of EC1 in classical cadherins was responsible for the homophilic interaction. It is now evident that the HAV sequence is specific to type I cadherins and that the corresponding sequences in type II cadherins and other members of the cadherin superfamily are different [Mahoney et al., 1991; Ranscht and Dours-Zimmermann, 1991; Suzuki et al., 1991; Sano et al., 1993; Tanihara et al., 1994a]. As described above, it is very difficult to believe that the HAV sequence is mainly responsible for the homophilic cell adhesion activity of classical cadherin. Indeed, T-cadherin, cadherin-11, protocadherin-1 and -2, and the *Drosophila* classical cadherin-related protein have no HAV sequence; but all these proteins were reported to have cell adhesion activity. Thus, the HAV sequence is not absolutely necessary for the cell adhesion activity. However, it is still possible that the HAV sequence constitutes the core structure of the binding site in conjunction with other sequences and plays an important role in strong cell-cell adhesion activity of classical cadherins. It would be interesting to know the effect of substitution of Histidine in the HAV sequence of E-cadherin with glutamine, since the corresponding sequence in various type II cadherins is QAV and many type II cadherins do not show strong cell-adhesion activity (see next section). It is noteworthy that the sequence of this region in EC1s is variable among different cadherins, although the corresponding sequences in other repeats are relatively similar, especially in non-classical cadherins. This result implies that the sequences in this region of EC1s are directly involved in the specific homophilic or heterophilic interaction. We have evidence that the EC1 of protocadherin-2 contains the cell-binding site [Obata et al., unpublished observation].

Another important prediction from this model is that cell-binding activity is conformation dependent; thus it is very sensitive to various structural perturbations.

(b) Ca^{2+}-binding site. One surprise of the structure model is that the Ca^{2+}-binding site is not directly linked to the putative homophilic binding site. The major role of Ca^{2+} seems to be a structural one, and the effect of Ca^{2+} on the binding activity is indirect, probably through the formation and maintenance of active conformation, although cadherins were initially identified as Ca^{2+}-requiring cell adhesion molecules. Several well conserved short amino acid sequences in the cadherin repeats were known from the beginning, and some of these sequences have been postulated as a Ca^{2+}-binding site [Ozawa et al., 1990]. According to the structure model, most of these sequences come close in a small region in the model and form a possible Ca^{2+}-binding site. Interestingly, the binding site is formed between the two adjacent cadherin repeats.

It is a well known property of classical cadherins that they are protected from trypsin digestion with Ca^{2+} [Takeichi, 1977]. Indeed, this property has played a key role during the initial identification of classical cadherins and the cell adhesion assay [Hyafil et al., 1981; Nagafuchi et al., 1987]. One possible explanation of this property could be that binding of Ca^{2+} to its binding site protects the trypsin-sensitive site by physically covering and/or by masking the sensitive sites through the conformation change induced by Ca^{2+}. Several experiments have indicated the conformation change of extracellular domain is induced by Ca^{2+} binding. However, recent experiments with type II cadherins, desmoglein 3, and protocadherins showed that these proteins were not effectively protected from trypsin digestion with Ca^{2+}. Thus, Ca^{2+} protection is not a general property of cadherins. It seems to be specific to type I cadherins and possibly to some other members of the cadherin superfamily.

Cell Adhesion and Other Activities of Cadherins

(a) Cell adhesion activity of classical cadherins. Since cell adhesion activity is the central issue for cadherin research, this activity has been studied extensively. The results overwhelm-

ingly indicate that type I cadherins have highly specific Ca^{2+}-dependent homophilic cell-cell adhesion activity. However, there is no direct in vitro assay system of the adhesion activity of cadherins and cadherin-related proteins, and no one has ever demonstrated the binding activity of the purified proteins in vitro. The evidence so far is from the cell aggregation assay of transfectants and from antibody inhibition experiments. This status is in marked contrast to other cell adhesion systems, such as integrins and N-CAM. Furthermore, various controversial results have been obtained regarding the cell adhesion activity of cadherins as described below.

Most of the extracellular domain of E-cadherin is released from the cell membrane with relatively high concentrations of trypsin in the presence of Ca^{2+}. The resultant fragment is stable, but there have been no published reports that this fragment forms a dimer. If E-cadherin has strong homophilic binding activity, it is natural to predict that the fragment has at least weak adhesiveness and forms dimers. Or if we put the purified type I cadherins into liposomes, again we could expect some binding activity; but no one has reported such data. Of course, one may argue that classical cadherins require cytoskeletal proteins for the strong cell adhesion activity; but if this were true, how could the cytoplasmic domain-deleted E-cadherin show the dominant negative effect [Levine et al., 1994]? Taken at face value, this result indicates that E-cadherin has homophilic interaction without the cytoplasmic domain, or it may have heterophilic interaction with other molecules as I will discuss later. These results suggest that the cell adhesion activity of type I cadherins may not be the simple process many investigators believe; rather it may be a complex process requiring additional unknown accessory molecules. Although it is highly unlikely, one possibility is that the apparent homophilic cell adhesion activity is not directly mediated by the type I cadherins, but is the secondary effect of cadherin function. In any event, it is unusual that the homophilic binding activity cannot be shown using purified proteins in vitro, and an in vitro assay system of the binding activity is definitely needed to delineate this controversial but very interesting activity of type I cadherins.

It is known that the interaction between the cytoskeletal proteins and the cytoplasmic domains of type I cadherins is essential for the cell adhesion activity [Ozawa et al., 1990]. The association of α-catenin to the cytoplasmic domains is especially important for the activity [Hirano et al., 1992]. Indeed, E-cadherin with α-catenin sequence instead of its original cytoplasmic domain sequence shows stable cell adhesion activity, indicating that α-catenin alone without other catenins can support the cell adhesion activity. However, the association between the α-catenin and the cytoplasmic domain was reported to be indirect, probably through β-catenin. This may explain the finding that a β-catenin mutant did not support the cell adhesion activity of E-cadherin. An interesting question is how the weak association can support the strong cell adhesion activity. Furthermore, Kintner [1992] showed that the recombinant N-cadherin that lacks a part of the extracellular domain showed a dominant negative effect, probably through competing catenins with native cadherins. Interestingly, the dominant negative effect was observed even for the mutated N-cadherin that lacks the C-terminal catenin-binding region. Taken at face value, this indicates that one or more additional essential molecules interact with the N-terminal region of the cytoplasmic domain, suggesting that the cell adhesion activity of cadherins is more complicated than is generally thought.

Another potentially serious problem is that the interaction of these proteins has mostly been examined by immunoprecipitation after detergent solubilization. It is generally thought that type I cadherins are linked to actin-based cytoskeleton and the interaction is so strong that the complex is not easily solubilized with detergent treatment. If this view of active cadherins is correct, the current concept was drawn from the examination of minor and possibly inactive cadherin forms in the cells. There is no doubt that the results obtained by immunoprecipitation experiments have clarified many of the essential properties of cadherin interaction inside the cells; however, this may not be the whole story and this may be one reason why no one has been able to show the clear image of active state cadherins. Similar problems exist in desmosomal cadherins research. A new approach may be needed to solve this problem.

Cell adhesion activity of the type II cadherins is more elusive. We carried out the conventional cell aggregation assay of two type II cadherins, cadherin-5 and cadherin-8, using the L cell transfectants, but we were unable to show any significant cell aggregation activity for them, although the expressed cadherins were localized in a Ca^{2+}-dependent manner, mainly at cell-cell contact

sites [Tanihara et al., 1994b; Kido et al., unpublished observation]. Immunoprecipitation of cadherin-5 and cadherin-8 showed that these proteins associated with catenins, although the association with α-catenin was relatively weak. Furthermore, the chimeric cadherin-4 with the cytoplasmic domain of cadherin-5 showed cell aggregation activity in a conventional assay system, whereas the chimeric cadherin-5 with cadherin-4 cytoplasmic domain did not, despite the fact that it associates well with α-catenin as the wild type cadherin-4 [Tanihara et al., 1994b]. Taken together, cadherin-5 and cadherin-8 do not seem to have strong cell-cell adhesion activity. However, we think that they probably have some type of homophilic or heterophilic interaction. First, the expressed cadherin-4 and cadherin-5 are mainly localized at cell-cell contact sites. Second, cadherin-5 transfectants form aggregates when they subjected to the static aggregation assay as described for a receptor type-protein tyrosine phosphatase [Gebbink et al., 1993]. On the other hand, cadherin-11, another type II cadherin, has been reported to show Ca^{2+}-dependent cell adhesion activity, but the details of the properties have not been published. As described previously, type II cadherins have unique sequence features. Whether type II cadherins as a group show distinctive adhesion properties from those of type I cadherins is an interesting question.

(b) Cell adhesion activity of protocadherins. Protocadherin-1 and protocadherin-2 transfectants show very weak cell aggregation activity in the conventional assay system using L cell transfectants [Sano et al., 1993]. To further study the details of the cell adhesion activity, we made a chimeric construct of protocadherin-2, substituting E-cadherin cytoplasmic sequence for the original cytoplasmic sequence. The resultant transfectants showed stronger cell aggregation activity that was Ca^{2+}-dependent and homophilic in nature. Furthermore, it showed specificity of cell aggregation. These findings clearly indicate that the extracellular domain of protocaderin-2 and possibly other protocadherins are capable of specific Ca^{2+}-dependent homophilic interaction. This is consistent with the recent report of Oda et al. [1994] that a protocadherin-related protein from *Drosophila* containing classical cadherin-like cytoplasmic sequence showed adhesion properties similar to those of classical cadherins. However, it is unclear whether protocadherin-2 functions as a typical cell-cell adhesion protein in vivo. Furthermore, various protocadherins appear to have different cell adhesion properties. We were unable to show any significant cell adhesion activity for the third protocadherin, protocadherin-3, which has a novel cytoplasmic domain. Nevertheless, we think that protocadherin-3 is involved in some type of homophilic or heterophilic interaction activity, since the expressed protein was localized at the cell-cell contact sites.

Protocadherins appear to interact with unique cytoplasmic proteins as predicted from the characteristic cytoplasmic domain sequences. The distinctive adhesion properties may reflect the biological role of protocadherins. The further study of these proteins may provide an *insight* into the biological functions in a Ca^{2+}-dependent manner.

(c) Cell adhesion activity of desmosomal cadherins. The cell adhesion activity of desmosomal cadherins has not been studied rigorously. Desmosomal cadherins appear to show no significant cell aggregation activity in L cell transfectants, although no reports have been published. Recently however, a study on the cell adhesion activity of desmoglein 3 using a chimeric protein with E-cadherin cytoplasmic domain has been published. The chimeric protein had weak cell aggregation activity which showed specificity. Again, this result is very difficult to understand if desmoglein 3 really has strong cell adhesion activity as in the type I classical cadherins. Interestingly though, E-cadherin containing the desmoglein cytoplasmic domain sequence instead of the original sequence showed strong cell adhesion activity. It is very difficult to explain these results. In any case, the cell adhesion activity of desmosomal cadherins does not seem to be as simple as the activity of classical cadherins.

(d) Other activities. Two other activities were reported for two members of the cadherin superfamily. Matsunaga et al. [1988] reported that N-cadherin has neurite outgrowth promoting activity. Since the activity is inhibited by the monoclonal antibody that inhibits the cell adhesion activity, the active site seems to be the same as the homophilic binding site or is present nearby. Since many proteins have neurite promoting activity, the mechanism of the activity may not be simple, and recent studies suggest that a signal transduction process is involved [Williams et al., 1994].

Another interesting study was recently reported that a cadherin-related protein participates in peptide transport in intestine [Dantzig

et al., 1994]. Although the activity was inhibited by an antibody against the protein, details of the mechanism are unknown. It seems unlikely that the protein itself is the transporter, since this activity is entirely different from the cell adhesion activity and the protein structure does not indicate any transporter feature. It seems that the protein is not directly involved in the process, but may be a secondary effect.

Cadherins are generally thought to be homophilic cell adhesion proteins. However, Cepek et al. [1994] reported that E-cadherin showed a *heterophilic* interaction with an integrin, $\alpha E\beta 7$. The interaction appears to be weak and the binding site is not exactly the same as the homophilic cell adhesion site. It is unclear whether the interaction has an actual biological role, but this report showed the first evidence that cadherins are capable of heterophilic interaction. Moreover, the possible interaction between N-cadherin and FGF has been suggested, and indirect interaction between E-cadherin and EGF receptor have recently been reported.

Available data suggest that the mechanism of cell adhesion activity of the cadherin superfamily members is not as simple as was thought a couple of years ago. Although many members appear to be capable of Ca^{2+}-dependent homophilic interaction, various circumstantial evidence suggests that different cadherins also have heterophilic interaction and other activities. It may be time to seriously consider these possibilities, since the results would profound impact further study of the cadherin superfamily, especially study of the biological functions.

Possible Biological Functions of Cadherins

Classical cadherins were identified as Ca^{2+}-dependent cell-cell adhesion molecules, and indeed recent gene targeting experiments with E-cadherin confirmed that E-cadherin is essential for cell-cell adhesion and that the disruption of the gene is lethal [Larue et al., 1994]. The major function of E-cadherin is to mediate the specific cell-cell adhesion and to play a pivotal role in the formation and maintenance of many epithelial tissues. As predicted from this, dysfunction of E-cadherin may be responsible, at least in part, for the metastatic capability of some malignant cells, and a vast number of reports supporting this hypothesis have been published in the last couple of years [for review see Birchmeier et al., 1995]. Desmosomal cadherins appear to have similar cell adhesion function and to act as major cell adhesion proteins in keratinocytes. An autoimmune disease is known to be desmosomal cadherin-related [Amagai, 1995].

Recent studies have revealed that various cadherins are expressed in a variety of multicellular organisms and most cell types express at least one type of cadherin. Furthermore, their expressions are developmentally regulated and show specific patterns. Cadherins appear to play an important role(s). However, the biological function of cadherins has become more elusive as the study has progressed.

Various functions that require cell adhesion activity to some extent have been suggested for different cadherins, such as cell layer segregation, axon guidance, and cell differentiation. Most of these suggested functions are related to morphogenesis. It is well documented that the expression of type I cadherins is developmentally regulated and that the expression pattern correlates well with morphogenetic events during embryogenesis. Indeed, the ectopic expression of classical cadherins in *Xenopus* has a profound effect on morphogenesis. So-called address code hypothesis that adhesion molecules function as the major determinant of tissue formation has been used to explain the role of adhesion molecules in embryogenesis. More generally, Edelman [1988] proposed a hypothesis of adhesion molecules as morphoregulatory molecules. However, the expression patterns of these proteins do not perfectly match the morphogenic processes when this is examined closely. More importantly, other than a few cases there is little data that directly connects the morphological changes and the expression of classical cadherins. Clearly, these two events need to be linked if they are really connected.

We now know that the properties of different members of the cadherin superfamily vary significantly as described above. Clearly, some of these results cannot be explained by a simple cell adhesion protein model and require a new explanation. First, some members show little or no cell adhesion activity and/or diffuse subcellular localization. Second, the expression of some cadherins is higher in adult than in fetus. Third, even the cell adhesion activity of classical cadherins reveals puzzling properties to be solved. Fourth, the inhibition of classical cadherin function has a profound effect on cell physiology [Kintner, 1992; Levine et al., 1994]. Considering the available information as described in this essay, many cadherins appear to participate in

more general cell-cell interaction processes than simple cell-cell adhesion. Recently, various cell adhesion proteins have been reported to participate in different signal transduction pathways. For example, some integrins are involved in apoptosis and growth factor signaling in addition to cell-extracellular matrix adhesion [Vuori and Ruoslahti, 1994; Brooks et al., 1994]. Similarly, some members of the cadherin superfamily, at least, may be involved in some signal transduction processes. It may be that some classical cadherins play a major regulatory role in junctional complex formation during which classical cadherins transmit regulatory signals and coordinate the process [Nelson, 1992]. A more complex possibility is that cadherins may be involved in more general signal transduction pathways. For example, classical cadherins may also participate in the *wnt* signal transduction pathway, since β-catenin is also a component involved in the signal transduction cascade [Peifer and Wieschaus, 1990; Heasman et al., 1994]. Doherty and his colleagues [1994] have already reported that the neurite outgrowth promoting activity of N-cadherin requires a signal transduction process.

It appears quite reasonable to assume that some members of the cadherin superfamily act as typical cell adhesion proteins, but others act as the mediators of more general cell-cell interaction in which the major activity is not cell adhesion but signal transduction. Of course, some members may have both functions.

CONCLUSION

We now know that the structure and function of cadherins is not as simple as was thought a couple of years ago; cadherins appear to have complex cell adhesion activity and possibly other interaction activity. Actual biological functions of many cadherins are poorly understood. Even some of the type I cadherins appear to require further study of their actual roles. How should these issues be addressed? A genetic approach, such as the ectopic expression and knockout of these genes, domain swapping, and in vitro mutagenesis, is one possibility. Another major area of study might be the interaction between the cytoplasmic domains of cadherins and the cytoplasmic proteins. A big question is what signals are cadherins able to transmit within cells? The answer will provide valuable information about their functions. A similar question also remains to be solved for the extracellular domains: what molecules can interact with the extracellular domains of cadherins?

Before ending this paper, I would like to emphasize the importance of further biochemical studies to address these issues. As described previously, the research into the cadherin superfamily lacks some fundamental biochemical studies in comparison to the studies of other adhesion molecules. Molecular biological studies revolutionized this field as well as other fields and should continue to be a powerful tool to study complex biological processes. However, we cannot solve many basic biological problems without in vitro biochemical experiments.

We can fairly say that an era of cadherin research ended with the publication of the structure model of cadherin repeats. Clearly, this is the turning point of cadherin research in many ways. We now have a clear structure model of cadherins, and a new concept of their functions is emerging. I believe study on the cadherin superfamily is entering a new era with further surprises waiting for us.

ACKNOWLEDGMENTS

The author thanks Drs. Miller, McClay, and Pettitt for providing unpublished data. The author also thanks Dr. J. Usukura for his comments and Ms. S. Clarke for editorial assistance. This research was supported in part by National Institute of Health grant NS32456 and by a grant from ICOS Corporation.

ADDED IN PROOF

At a recent Keystone symposium Gumbiner's group provided convincing evidence using an in vitro system that a classical cadherin actually forms dimer. On the other hand, there is a growing consensus, based on the recent advances in cadherin research, that clustering of cadherins plays a major role in the strong cell-cell adhesion activity. Our results of chimeric cadherin experiments are consistent with this notion. Indeed, Shapiro et al. [1995] have proposed even a zipper model (Nature 374:327–337).

REFERENCES

Amagai M (1995): Adhesion Molecules. I: Keratinocyte-keratinocyte interactions; cadherins and pemphigus. J Invest Dermatol 104:146–152.

Birchmeier WE (1995): Cadherin as a tumor (invasion) suppressor gene. Bioessays 17:97–99.

Blaschuk OW, Pouliot Y, Holland, PC (1990): Identification of a conserved region common to cadherins and influenza strain A hemagglutinins. J Mol Biol 211:679–682.

Brooks PC, Montgomery AMP, Rosenfeld M, Reisfeld, RA, Hu T, Klier G, Cheresh DA (1994): Integrin $\alpha_{v\beta3}$ antagonists promote tumor regression by inducing apoptosis of angiogenic blood vessels. Cell 79:1157–1164.

Cepek KL, Shaw SK, Parker CM, Russell GJ, Morrow JS, Rimm DL, Brenner MB (1994): Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the $\alpha^E\beta_7$ integrin. Nature 372:190–193.

Dantzig AH, Hoskins JA, Tabas LB, Bright S, Shepard RL, Jenkins IL, Duckworth DC, Sportsman JR, Mackensen D, Rosteck PR Jr., Skatrud PL (1994): Association of intestinal peptide transport with a protein related to the cadherin superfamily. Science 264:430–433.

Edelman GM (1988): Topobiology: An Introduction to Molecular Embryology. New York: Basic Books.

Gebbink MFBG, Zondag GCM, Wubbolts RW, Beijersbergen RL, Van Etten I, Moolenaar WH (1993): Cell-cell adhesion mediated by a receptor-like protein tyrosine phosphatase. J Biol Chem 268:16101–16104.

Geiger B, Ayalon O (1992): Cadherins. Annu Rev Cell Biol 8:307–332.

Hatta K, Nose A, Nagafuchi A, Takeichi M (1988): Cloning and expression of DNA encoding a neural calcium-dependent cell adhesion molecule: its identity in the cadherin gene family. J Cell Biol 106:873–881.

Heasman J, Crawford A, Goldstone K, Garner-Hamrick P, Gumbiner B, McCrea P, Kintner C, Noro CY, Wylie C (1994): Overexpression of cadherins and underexpression of β-catenin inhibit dorsal mesoderm induction in early *Xenopus* embryos. Cell 79:791–803.

Hirano S, Kimoto N, Shimoyama Y, Hirohashi S, Takeichi M (1992): Identification of a neural α-catenin as a key regulator of cadherin function and Multicellular organization. Cell 70:293–301.

Hyafil F, Babinet C, Jacob F (1981): Cell-cell interactions in early embryogenesis: A molecular approach to the role of calcium. Cell 26:447–454.

Hynes RO, Lander AD (1992): Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. Cell 68:303–322.

Kintner C (1992): Regulation of embryonic cell adhesion by the cadherin cytoplasmic domain. Cell 69:225–236.

Koch PJ, Walsh MJ, Schmelz M, Goldschmidt MD, Zimbelmann R, Franke WW (1990): Identification of desmoglein, a constitutive desmosomal glycoprotein, as a member of the cadherin family of cell adhesion molecules. Eur J Cell Biol 3:1–12.

Larue L, Ohsugi M, Hirchenhain J, Kemler R (1994): E-cadherin null mutant embryos fail to form a trophectoderm epithelium. Proc Natl Acad Sci USA 91:8263–8267.

Levine E, Lee CH, Kintner C, Gumbiner BM (1994): Selective disruption of E-cadherin function in early *Xenopus* embryos by a dominant negative mutant. Development 120:901–909.

Mahoney PA, Weber U, Onofrechuk P, Biessmann H, Bryant PJ, Goodman CS (1991): The *fat* tumor suppressor gene in Drosophila encodes a novel member of the cadherin gene superfamily. Cell 67:853–868.

Matsunaga M, Hatta K, Nagafuchi A, Takeichi M (1988): Guidance of optic nerve fibres by N-cadherin adhesion molecules. Nature 334:62–64.

Miyatani S, Copeland NG, Gilbert DJ, Jenkins NA, Takeichi M (1992): Genomic structure and chromosomal mapping of the mouse N-cadherin gene. Proc Natl Acad Sci USA 89:8443–8447.

Nagafuchi A, Shirayoshi Y, Okazaki K, Yasuda K, Takeichi M (1987): Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA. Nature 329: 341–343.

Nelson WJ (1992): Regulation of surface polarity from bacteria to mammals. Science 258:948–955.

Nose A, Tsuji K, Takeichi M (1990): Localization of specificity determining sites in cadherin cell adhesion molecules. Cell 61:147–155.

Oda H, Uemura T, Harada Y, Iwai Y, Takeichi M (1994): A *Drosophila* homolog of cadherin associated with armadillo and essential for embryonic cell-cell adhesion. Dev Biol 165:716–726.

Overduin M, Harvey TS, Bagby S, Tong KI, Yau P, Takeichi M, Ikura M (1995): Solution structure of the epithelial cadherin domain responsible for selective cell adhesion. Science 267:386–389.

Ozawa M, Baribault H, Kemler R (1989): The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. EMBO J 8:1711–1717.

Ozawa M, Engel J, Kemler R (1990): Single amino acid substitutions in one Ca^{2+} binding site of uvomorulin abolish the adhesive function. Cell 63:1033–1038.

Ozawa M, Hoschuetzky H, Herrenknecht K, Kemler R (1990): A possible new adhesive site in the cell-adhesion molecule uvomorulin. Mech Dev 33:49–56.

Peifer M, Wieschaus E (1990): The segment polarity gene *armadillo* encodes a functionally modular protein that is the Drosophila homolog of human plakoglobin. Cell 63: 1167–1179.

Ranscht B, Dours-Zimmermann MT (1991): T-cadherin, a novel cadherin cell adhesion molecule in the nervous system lacks the conserved cytoplasmic region. Neuron 7:391–402.

Sano K, Tanihara H, Heimark RL, Obata S, Davidson M, St. John T, Taketani S, Suzuki S (1993): Protocadherins: a large family of cadherin-related molecules in central nervous system. EMBO J 12:2249–2256.

Suzuki S, Sano K, Tanihara H (1991): Diversity of the cadherin family: evidence for eight new cadherins in nervous tissue. Cell Reg 2:261–170.

Takeichi M (1977): Functional correlation between cell adhesive properties and some cell surface proteins. J Cell Biol 75:464–474.

Takeichi M (1991): Cadherin cell adhesion receptors as a morphogenetic regulator. Science 251:1451–1455.

Tanihara H, Sano K, Heimark RL, St. John T, Suzuki S (1994a): Cloning of five human cadherins clarifies characteristic features of cadherin extracellular domain and provides further evidence for two structurally different types of cadheris. Cell Adhesion and Communication 2: 15–26.

Tanihara H, Kido M, Obata S, Heimark RL, Davidson M, St. John T, Suzuki S (1994b): Characterization of cadherin-4 and cadherin-5 reveals new aspects of cadherins. J Cell Sci 107:1697–1704.

Townes PL, Hotfreter J (1955): Directed movements and selective adhesion of embryonic ampibian cells. J Exp Zool 128:53–120.

Vuori K, Ruoslahti E (1994): Association of insulin receptor substrate-1 with integrins. Science 266:1576–1578.

Williams EJ, Furness J, Walsh FS, Doherty P (1994): Activation of the FGF receptor underlies neurite outgrowth stimulated by L1, N-CAM, and N-cadherin. Neuron 13: 583–594.