The EMBO Journal vol.12 no.6 pp.2249–2256, 1993

# Protocadherins: a large family of cadherin-related molecules in central nervous system

**Kenji Sano[1], Hidenobu Tanihara[1], Ronald L.Heimark[3], Shuichi Obata[1], Mari Davidson[3], Tom St.John[3], Shigeru Taketani[4] and Shintaro Suzuki[1,2,5]**

[1]Doheny Eye Institute and [2]Departments of Ophthalmology and Microbiology, University of Southern California School of Medicine, Los Angeles, CA 90033, [3]ICOS Corporation, 22021 20th Avenue, SE, Bothell, WA 98021, USA and [4]Department of Hygiene, Kansai Medical University, Fumizono-cho, Moriguchi 570, Japan

[5]Corresponding author

Communicated by J.P.Thiery

**Using the polymerase chain reaction, we have isolated numerous rat and human cDNAs of which the deduced amino acid sequences are highly homologous to the sequences of the extracellular domain of cadherins. The entire putative coding sequences for two human proteins defined by two of these cDNAs have been determined. The overall structure of these molecules is very similar to that of classic cadherins, but they have some unique features. The extracellular domains are composed of six or seven subdomains that are very similar to those of cadherins, but have characteristic properties. The cytoplasmic domains, on the other hand, have no significant homology with those of classic cadherins. Since various cDNAs with almost identical features were obtained also from *Xenopus*, *Drosophila* and *Caenorhabditis elegans*, it appears that similar molecules are expressed in a variety of organisms. We have tentatively named these proteins protocadherins. They are highly expressed in brain and their expression appears to be developmentally regulated. The proteins expressed from the two full-length cDNAs in L cells were ~170 or 150 kDa in size, and were localized mainly at cell–cell contact sites. Moreover, the transfectants showed cell adhesion activity.**

*Key words*: cadherins/central nervous system/extracellular domain/protocadherins/uvomorulin

## Introduction

Cadherins are a group of integral membrane proteins with a mol. wt of ~130 kDa that mediate calcium-dependent cell–cell adhesion (for a review see Kemler *et al.*, 1989; Takeichi, 1991). The molecules are composed of a large extracellular domain at the N-terminal side and a relatively small cytoplasmic domain at the C-terminal side; these domains are connected by a transmembrane sequence. The extracellular domain consists primarily of five subdomains with a cadherin-specific motif (cadherin motif) and contains the cell–cell interaction site that determines the specificity of cadherins. The cytoplasmic domain is thought to regulate cadherin function through interactions with cytoskeletal proteins.

Until recently, only three cadherins, E-, N- and P-cadherin, had been identified and well characterized (for a review see Takeichi, 1991). These molecules are also referred to by other names, such as uvomorulins (Kemler *et al.*, 1977), L-CAM (Bertolotti *et al.*, 1980) and A-CAM (Volk and Geiger, 1986). Recent studies, however, have uncovered many new cadherins, and the cadherin family has turned out to be a large one (Donalies *et al.*, 1991; Ginsberg *et al.*, 1991; Inuzuka *et al.*, 1991; Napolitano *et al.*, 1991; Suzuki *et al.*, 1991). Furthermore, cadherin-related molecules such as desmoglein, desmocollins and T-cadherin have been described (Goodwin *et al.*, 1990; Holton *et al.*, 1990; Koch *et al.*, 1990; Ranscht and Dours-Zimmermann, 1991). Their extracellular domains share common properties with those of typical cadherins; however, desmoglein and desmocollins have a characteristic cytoplasmic domain and T-cadherin lacks the cytoplasmic domain. In addition, a tumor suppressor gene of *Drosophila*, called *fat*, has recently been shown to encode a cadherin-related protein (Mahoney *et al.*, 1991). It comprises 34 tandem repeats of a cadherin motif, followed by four EGF-like repeats, a transmembrane domain and a novel cytoplasmic domain. The repeated extracellular subdomains of *fat* share many properties with those of the typical cadherins. These findings suggest that cadherins and cadherin-related molecules are expressed in various tissues of different organisms, and constitute a large superfamily characterized by the presence of the cadherin motif.

Cadherins are thought to play a major role in cell–cell adhesion and thus are involved in various biological processes, such as morphogenesis and maintenance of tissue structure. In support of these roles for cadherins, it has been shown that the expression of cadherins is spatially and temporally regulated during development, and that each cadherin is expressed by specific cells and tissues at a specific stage (for a review see Takeichi, 1991). Moreover, ectopic cadherin expression or inhibition of the normal expression hinders the formation of normal tissue structure (Detrick *et al.*, 1990; Fujimori *et al.*, 1990; Kintner, 1992). Autoantibodies from patients with pemphigus vulgaris, an autoimmune skin disease characterized by blister formation caused by loss of cell adhesion, react with a cadherin-related molecule, offering direct support for an adhesion function of cadherins *in vivo* (Amagai *et al.*, 1991). On the other hand, mammalian E-cadherin and *Drosophila fat* function as a tumor invasion suppressor or a tumor suppressor, suggesting that cadherins and cadherin-related cell adhesion proteins directly or indirectly regulate cell behavior or cell growth (Chen and Obrink, 1991; Frixen *et al.*, 1991; Mahoney *et al.*, 1991; Vleminckx *et al.*, 1991).

To study the function of cadherins and cadherin-related molecules, we initiated an investigation into the possibility that other cadherins and cadherin-related molecules, as yet unidentified, might participate in various cell–cell interactions. Using a polymerase chain reaction (PCR)

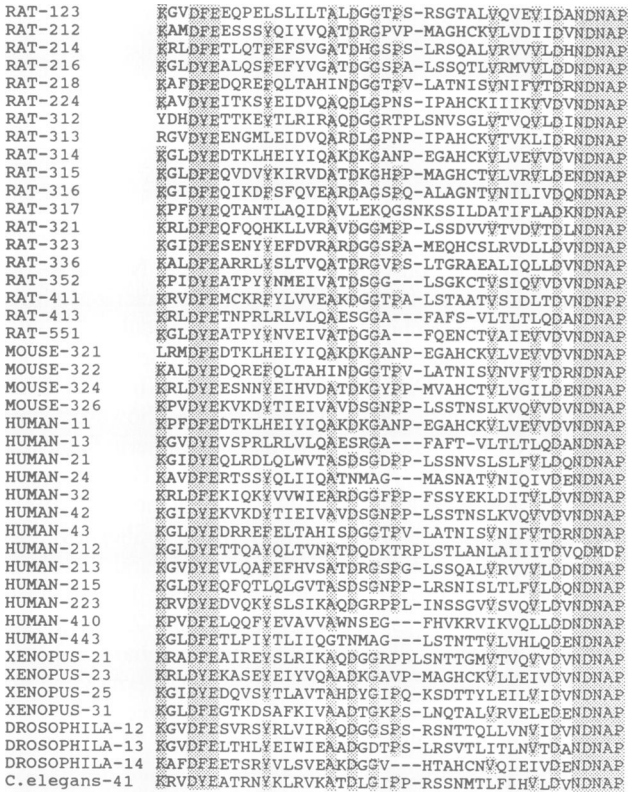


**Fig. 1.** Deduced amino acid sequences of cDNA clones isolated by PCR. All of the clones were isolated from cDNA preparations made from the nervous tissues or derivative cells from rat, mouse, human and *Xenopus*, except for the *C.elegans* and *Drosophila* clones, which were isolated from preparations made from the whole body. The highly conserved amino acids are shaded. The sequences corresponding to the used primers may contain PCR artifacts.

method, we have previously identified several new cadherins (Suzuki *et al.*, 1991). Here we report the isolation of a large number of cDNAs of which deduced amino acid sequences are very similar to, but distinct from, those of the known cadherins and cadherin-related proteins.

## Results

### *Isolation of cDNAs for new cadherin-related molecules*

A PCR method was used to isolate cDNAs that encode the cadherin extracellular domain motif. Two conserved amino acid sequences, one from the middle of the third extracellular subdomain (EC-3) and the other from the C-terminal of EC-4, were chosen by comparing various cadherin sequences (Gallin *et al.*, 1987; Nagafuchi *et al.*, 1987; Nose *et al.*, 1987; Ringwald *et al.*, 1987; Hatta *et al.*, 1988; Miyatani *et al.*, 1989; Shimoyama *et al.*, 1989) and the corresponding degenerate oligonucleotides were designed for use as PCR primers (for sequences see Materials and methods). PCR was carried out using these mixed oligonucleotides as primers and a cDNA preparation from rat brain as template. The resultant product showed two major bands, of ~450 and 130 bp in size, on agarose gel electrophoresis. The 450 bp band corresponds to the length between the two primer sites of cadherins, but the 130 bp band would not be predicted from any of the known cadherin sequences. Therefore, the 130 bp product was subcloned into the M13 vector, and ~50 clones were isolated and their DNA sequences were determined. The resultant sequences revealed ~20 different types of cDNA clones of which the deduced amino acid sequences are homologous to, but distinct from, the C-terminal regions of the extracellular domain repeats of the known cadherins (Figure 1). All of the cadherins reported thus far have characteristic, well-conserved, short amino acid sequences in the EC-3; DFE or DYE sequence is located at the middle region, and DXNEXPX or DXDEXPX sequence is at its end region (Hatta *et al.*, 1988), whereas the corresponding sequences of other subdomains, except for EC-5, are DRE and DXNDNXPX, respectively. The deduced amino acid sequences of the new clones have DFE or DYE at one end, but have DXNDNXPX instead of DXNEXPX or DXDEXPX at the other end. To our knowledge, this combination has not been reported in any other known cadherin. Moreover, the sizes of the new clones are very similar and AXDXG in the middle region is highly conserved among the deduced amino acid sequences.

cDNAs with similar features were also isolated from human, mouse, *Xenopus*, *Drosophila* and *Caenorhabditis elegans* (Figure 1). Most of these clones were isolated from brain, retina or neuronal and glial cell lines, and were not identified in most other tissues or their derived cell lines. Comparison of the sequences indicates significant similarity between sets of these clones. In particular, there were three sets that appear to be cross-species homologues: rat clone 218, mouse clone 322 and human clone 43; rat clone 314, mouse clone 321 and human clone 11; mouse clone 326 and human clone 42.

### *Deduced amino acid sequences of two new cadherin-related proteins*

To ascertain the structure of the new proteins defined by the cloned PCR products, two proteins, tentatively named pc42 and pc43, were randomly chosen, their cDNAs were obtained and the entire coding sequences were determined. A human brain cDNA library was screened with the short cDNAs obtained by PCR, and the resultant clones were sequenced. Several overlapping cDNA clones contained the entire putative coding sequences of two proteins. The nucleotide sequence data have been deposited in the EMBL/GenBank Data Library under the accession numbers L11370 and L11373. The overall structure of these proteins is similar to that of typical cadherins. Both cDNA sequences contain possible translation initiation sites that match the Kozak criteria (Kozak, 1984). The translated amino acid sequences start with typical signal sequences, but they lack the prosequences that are present in all other known cadherin precursors. There is an extracellular domain at the N-terminal side and a cytoplasmic domain at the C-terminal side; these domains are connected by a transmembrane sequence. The extracellular domains of pc42 and pc43 proteins are different in length. Pc42 contains seven and pc43 has six repeats of a motif that resembles closely the typical cadherin extracellular domain motif (Figure 2). The sizes of the cytoplasmic domains are similar to those of typical cadherins, but the sequences do not show any significant homology with those of known cadherins, cadherin-related molecules or any other sequences in the data bank.

Some of the clones contain unique sequences at the 3′ end region that correspond to the cytoplasmic domain or the C-terminal region of the extracellular domain, although the remaining 5′ side sequence is identical to the prototype sequences. The boundary sequences are consistent with the

**Fig. 2.** Alignment of deduced amino acid sequences of human pc42 and pc43, mouse N-cadherin and *Drosophila fat*. Extracellular domain repeats of human pc42 and pc43 and N-cadherin, *fat* EC-18 were aligned. Gaps were introduced to maximize the homology and the highly conserved amino acids were shaded. The amino acid residues indicated by capital letters in the motif are present in more than half of the repeats of various cadherins, human pc42 and pc43, and *Drosophila fat*. The amino acid residues described by small letters are relatively well conserved in human pc42 and pc43, and *Drosophila fat*. O in the 'Motif' line denotes aromatic amino acids.

consensus sequence of mRNA splicing (Cech, 1983), suggesting that these clones correspond to alternatively spliced mRNAs. (These nucleotide sequence data have been deposited in the EMBL/GenBank Data Library: accession numbers are L11369, L11371 and L11372.)

### *Comparison of extracellular domain amino acid sequences of pc42 and pc43 proteins and other sequences*

The deduced amino acid sequences of the extracellular subdomains of pc42 and pc43 proteins are aligned with those of N-cadherin EC-1−EC-5, and *Drosophila fat* EC-18, which represents a typical *fat* extracellular subdomain (Figure 2). Many amino acids characteristic of other cadherin extracellular domain repeats are conserved in the new sequences. The sequences DXD, DRE and DXNDNXPXF, two glycine residues and one glutamic acid residue, are the major conserved features among the protocadherins and classic cadherins. The new sequences, however, contain additional features. More amino acids at specific sites are conserved, such as DXDXGXN sequence near the N-terminal of the repeat and AXDXGXP sequence near the C-terminal (see Motif in Figure 2). Another noticeable property is that the lengths between the well-conserved sequences and amino acid residues described above are almost identical among the repeats, except for the regions near the C-terminal of EC-1, the middle of EC-2 and EC-4, and the C-terminal of the last repeat, which do not show significant homology with corresponding regions of other repeats (Figure 2). In contrast, each repeat of classic cadherins has characteristic features and size, with the exception of cadherin EC-2, which has some features similar to the newly found repeats. The extracellular domain repeats of *fat* and of molecules pc42 and pc43 share almost identical features.

The homologies among the extracellular subdomain sequences of pc42 and pc43 are not high (<40%), despite the fact that key amino acids of the cadherin motif are highly conserved. Similarly, the amino acid identity values between the extracellular subdomains of the new molecules, and N-cadherin EC-1−EC-5 and *Drosophila fat* EC-18 are mostly <40%. These identity values are comparable to the values between the subdomains of various cadherins (Gallin *et al.*, 1987; Nagafuchi *et al.*, 1987; Nose *et al.*, 1987; Ringwald *et al.*, 1987; Hatta *et al.*, 1988; Donalies *et al.*, 1991; Ginsberg *et al.*, 1991; Inuzuka *et al.*, 1991; Napolitano *et al.*, 1991; Suzuki *et al.*, 1991).

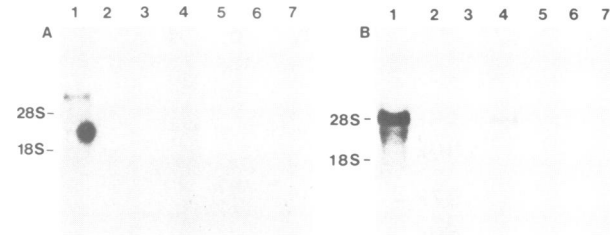


**Fig. 3.** Expression of mRNAs for pc42 and pc43 in rat tissues. Total mRNA preparations of rat tissues (10 μg/lane) were separated electrophoretically under denaturing conditions and transferred onto nitrocellulose filters. The resultant filters were hybridized with $^{32}$P-labeled mouse probe (mouse clone #326) corresponding to the EC-4 of pc42 **(A)** and rat probe (rat clone #218) corresponding to the EC-5 of pc43 **(B)**. Calf liver ribosomal RNAs (28S and 18S) were used as size markers. 1, brain; 2, heart; 3, liver; 4, lung; 5, skin; 6, kidney; 7, muscle.

The extracellular domains of molecules pc42 and pc43 have similar subdomain organization. As described above, the amino acid identity values are not as high as in the case of classic cadherins, but both proteins share regions that do not show significant homology with the typical cadherin motif near the C-terminal of EC-1, in the middle of EC-2 and EC-4, and at the C-terminal of the last repeat (Figure 2). One cysteine residue is located at a similar position in the middle of EC-4. These results suggest that pc42 and pc43 are more closely related to one another than they are to classic cadherins, despite the differences in subdomain number and cytoplasmic domain sequence. We suggest the name protocadherins for this family of cadherin-related proteins.

### *Expression of pc42 and pc43 mRNAs*

The expression of the mRNAs for pc42 and pc43 proteins was examined by Northern blot analysis. The mRNAs of both molecules were detected strongly in brain (Figure 3), an expected result based on the results of PCR. The pc42 probe detected a major band of 7 kb and a minor band of 4 kb in size (Figure 3A), possibly representing the products of alternative splicing. The pc43 probe, on the other hand, hybridized with a major band of 5 kb in size and with minor bands of smaller sizes (Figure 3B).

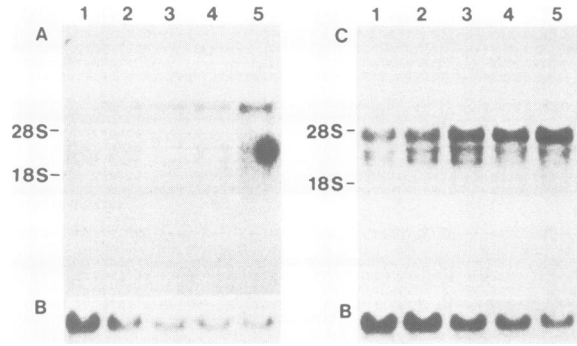


**Fig. 4.** Developmental change in mRNA expression of rat pc42 and pc43. Total RNA preparations were made from rat brains at embryonic days 17 (1) and 20 (2), neonatal days 5 (3) and 11 (4) and adult brain (5). Experimental procedures of Northern blot analysis were the same as described in Figure 3. β-Actin was used as an internal standard. **A**, pc42; **B**, β-actin; **C**, pc43.

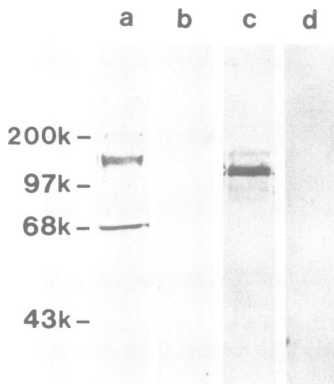


**Fig. 5.** Immunoblot analysis of pc42 and pc43 transfectants. The extracts of pc42 and pc43 transfectants were subjected to SDS–PAGE and then blotted electrophoretically to a PVDF membrane. The anti-pc42 antibody stained a band with ~170 kDa mol. wt in pc42 transfectant cells **(a)**, but not parental L cells **(b)**. The anti-pc43 monoclonal antibody specifically stained two adjacent bands of ~150 kDa in pc43-transfected cells **(c)**. The antibody did not stain the bands in parental L cells **(d)**.

Some neuronal and glial cell lines were found to express these protocadherin messages. SK-N-SH human neuroblastoma cells and U251 human glioma cells expressed the pc43 mRNA, and Neuro-2a mouse neuroblastoma cells expressed the pc42 mRNA.

To examine the developmental regulation of mRNA expression of these molecules, brain mRNA from rats at embryonic days 17 and 20, neonatal days 5 and 11, and from adult rats, was prepared and subjected to Northern blot analysis. As shown in Figure 4, the mRNA levels for pc42 and pc43 proteins increased during development, as compared with β-actin expression.

### *Characterization of pc42 and pc43 transfectants*

In order to determine some of the properties of the pc42 and pc43 proteins, their cDNAs were subcloned into the multicloning site of pRc/RSV expression vector and L cells were transfected with the resultant constructs. Stable transfectants were isolated by screening with G418. Subsequent experiments were carried out using two independent clones for each protocadherin and similar results were obtained for both clones. Northern blot analysis showed

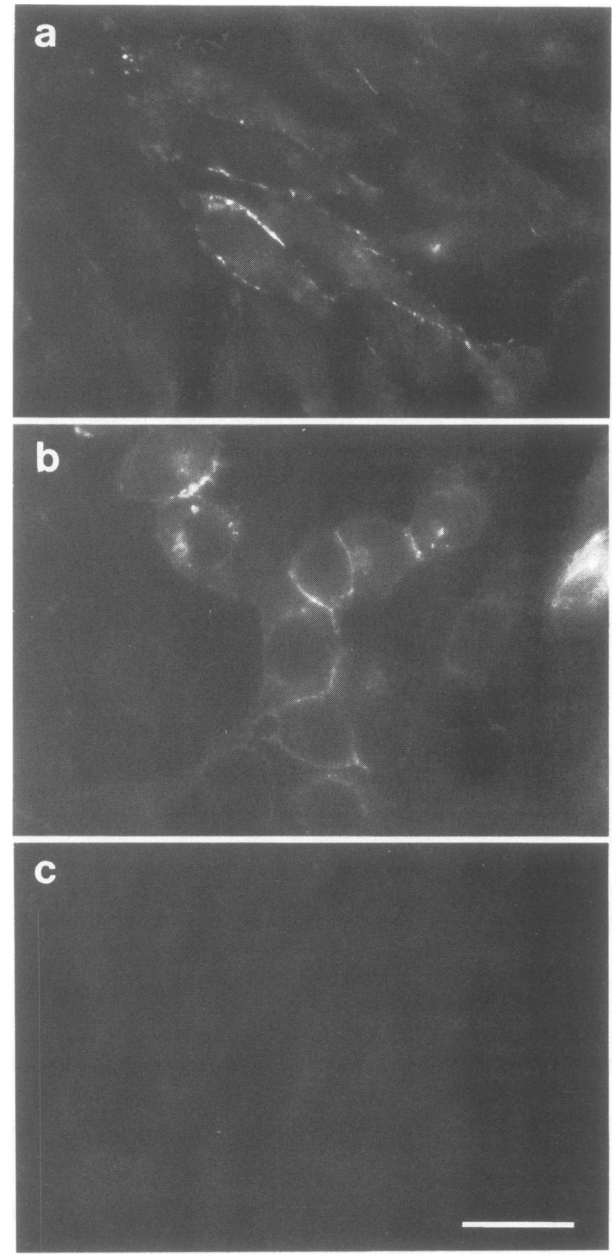


**Fig. 6.** Immunofluorescent staining of pc42 and pc43 transfectant cells. The transfectant cells grown on a cover-slip were fixed with cold methanol followed by paraformaldehyde and stained with anti-pc42 **(a)** and anti-pc43 **(b)** antibodies. The antibodies stained mainly cell–cell contact sites, but pre-immune antibody did not stain these sites **(c)**. Bar, 20 μm.

the expression of mRNAs of the expected sizes in the transfected cells (data not shown). On immunoblot analysis, an antibody against pc42 detected a specific 170 kDa band in the pc42 transfectant cells, while the antibody against pc43 specifically stained two adjacent bands with a mol. wt of ~150 kDa in the transfectants (Figure 5). A similar molecular weight for pc43 was obtained in SK-N-SH neuroblastoma cells (data not shown). These sizes are significantly larger than the molecular weights obtained from the deduced amino acid sequences. This discrepancy between predicted and actual molecular weights is common among various cadherins and is attributable to the glycosylation and/or cadherin-specific structural properties. The antibody

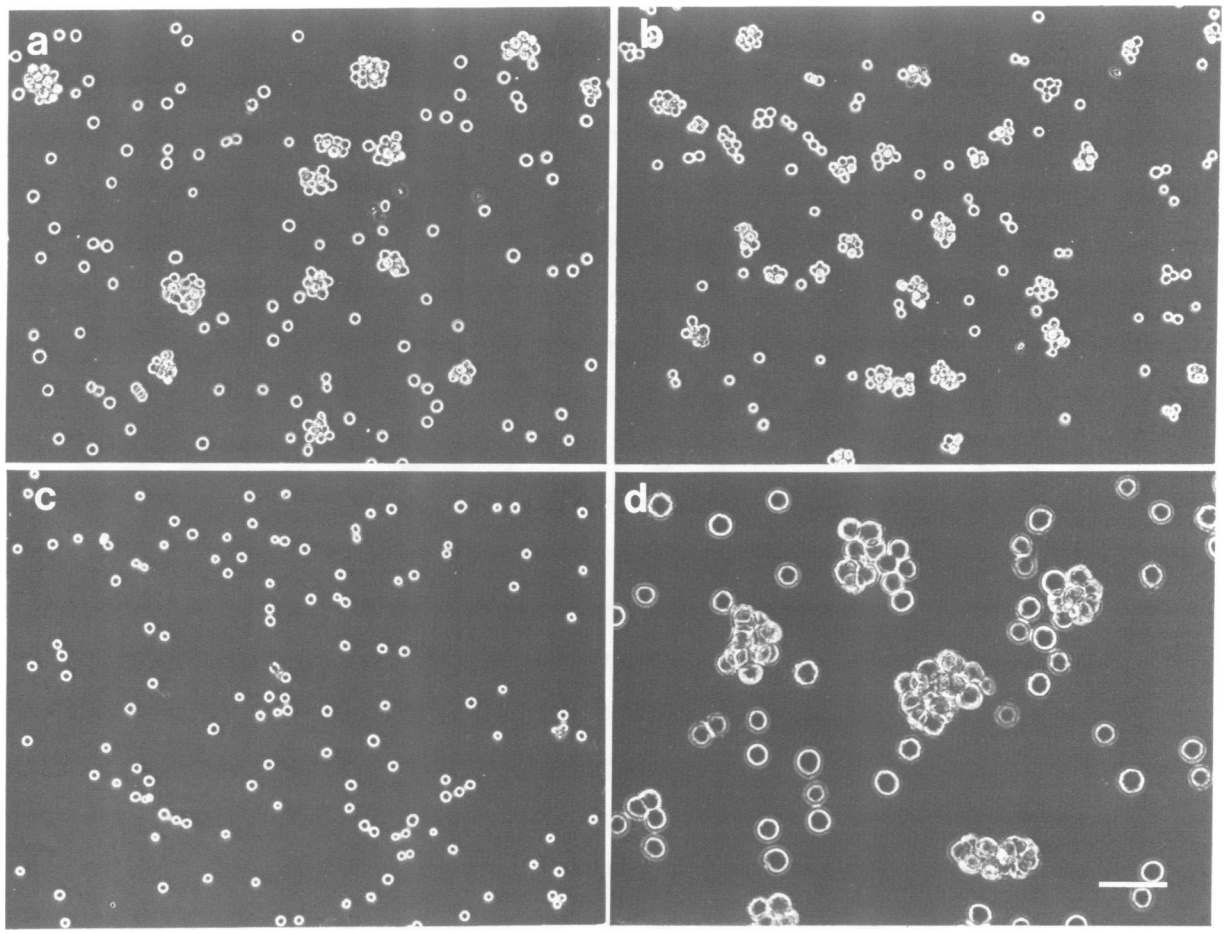


**Fig. 7.** Cell aggregation of pc42 and pc43 transfectant cells. Pc42 and pc43 transfectant cells were treated with 0.01% trypsin and 1 mM EGTA, and single cells were collected by centrifugation. After washing, the cells were incubated in the 1:1 mixture of DMEM and HBS containing 2 mM $CaCl_2$, 1% BSA and 20 μg/ml of deoxynucleotidase for up to 3 h on a rotary shaker. The pc42 and pc43 transfectant cells showed small aggregates after ~2.5 h incubation **(a, d)** and 1.5 h incubation **(b)**, respectively. Parental L cells did not form aggregates under the same conditions **(c)**. Bar, 100 μm (a–c) and 50 μm (d).

against pc42 also detected smaller bands, which may be the proteolytic degradation products.

The pc42 and pc43 proteins expressed in L cells were highly sensitive to proteolysis and were easily digested by a 30 min 0.01% trypsin treatment. The upper band of pc43 was more sensitive to treatment than the lower band. In contrast to the classic E-, N- and P-cadherins, however, pc42 and pc43 proteins were not protected from digestion in the presence of 1–5 mM $Ca^{2+}$. The pc43 protein of SK-N-SH cells showed a similar property.

The pc42 and pc43 transfectants were morphologically similar to the parental cells. In immunofluorescence microscopy, however, the antibodies against these proteins stained the periphery of the cells, primarily the cell–cell contact sites (Figure 6). Some cells showed the immunostaining of perinuclear areas which may correspond to Golgi apparatus. These staining results do not seem to be specific to the transfectants, since similar staining was obtained when human SK-N-SH neuroblastoma cells were stained with the anti-pc43 antibody. These antibodies did not stain the parent L cells and pre-immune sera did not stain the pc42 and pc43 transfectants.

Since the pc42 and pc43 sequences have the cadherin specific motif in their extracellular domains, the cell aggregation properties of these transfectant cells were examined. In this study, trypsin treatment in the presence of 1 mM EGTA was used to prepare single cells, since $Ca^{2+}$ did not protect pc42 and pc43 from trypsin digestion, and since this treatment easily yields single cells. The pc42 and pc43 transfectants did not show any significant cell aggregation activity during a short period of incubation (< 30 min). This is in contrast to what is seen with the classic cadherins in similar experiments (Nagafuchi *et al.*, 1987; Hatta *et al.*, 1988; Ozawa *et al.*, 1990). However, prolonged incubation (> 1–2 h) resulted in gradual formation of small aggregates (Figure 7). These activities were not observed when untransfected L cells or L cells transfected with pRc/RSV vector alone were used.

## Discussion

We have determined the entire coding sequences for two human cadherin-like proteins. The overall structure of these proteins is similar to that of typical cadherins: a relatively large extracellular domain consisting primarily of cadherin repeats is present at the N-terminal side, followed by a transmembrane sequence and a cytoplasmic domain. In contrast to the classic cadherins, the two new proteins have entirely distinct cytoplasmic domains. We propose that these proteins be designated protocadherins because of their unique structural features and the existence of similar proteins in distantly related species, as discussed below.

The protocadherin family is characterized by a distinct cadherin repeat in the extracellular domain. Our PCR-derived clones indicate that this type of cadherin repeat is present in various molecules from a variety of organisms, although we have not yet determined longer sequences for most of the proteins defined by the cDNAs from non-human species. In addition to the full-length human protocadherins, cDNA fragments encoding for this repeat were isolated from rat, mouse, *Xenopus*, *Drosophila* and *C.elegans*. Moreover, previously described *Drosophila fat* protein contains the same type of cadherin repeat as described here; one clone we isolated actually corresponds to *fat* EC-25. Therefore, the cadherin repeats characteristic of protocadherins are not specific to higher organisms, but are also present in many proteins from a variety of organisms.

It is conceivable that cadherin repeats are derived from the duplication and diversification of a primordial sequence, since the cadherin motif is highly conserved and widely distributed in various organisms, as described above. If this is the case, it is highly likely that the common sequence features found presently in the various proteins may be very similar to the sequence of the original cadherin motif. Therefore, we believe that the cadherin motif described here, which seems to be present in many proteins from a variety of organisms, retains many features of the primordial cadherin motif (see the Motif in Figure 2), hence the name protocadherins. It should be pointed out, however, that recent clonings of several cadherin genes have revealed that the genomic structures of EC-1−EC-5 are not the same (Hatta *et al.*, 1991; Ringwald *et al.*, 1991; Sorkin *et al.*, 1991). The genomic structure of various cadherin repeats may have changed extensively during evolution. Further studies on the genomic structure of these new cadherin-related molecules may clarify this issue.

We have already isolated ~30 different cDNA clones from human, rat and mouse. The proteins defined by the cDNAs appear to belong to the protocadherin family, judging from their sequences. Since we have not done a thorough screening to isolate similar clones, we expect that many clones with the same sequence properties will be found in the near future. We used a relatively unique sequence, DFE or DYE, for the synthesis of one PCR primer, but this sequence is clearly present in more than one repeat in some molecules. Therefore, the number of new molecules defined by the cDNAs is actually lower than the number of new clones.

The function of the protocadherins is unclear. Protocadherins are likely to serve a basic and important function, since many protocadherins are expressed in a variety of organisms, and since the size of the repeats and the key amino acids are conserved well during evolution.

One important issue is whether or not protocadherins are functional adhesion proteins. All of the cadherins and cadherin-related molecules reported thus far are thought to be involved in cell−cell interactions, mainly in cell−cell adhesion processes. Transfection experiments using full-length cDNA of pc42 and pc43 demonstrated that the expressed proteins could be found at the cell periphery, mainly at the cell−cell contact sites. Moreover, the pc42 and pc43 transfectants showed cell aggregation activity. The cell adhesion activities were specific to the transfectants. Parental L cells and transfectants with vector only did not show significant cell adhesion activity under the conditions that we used. Furthermore, transfectants of cadherin-5 and cadherin-8, which are new typical cadherins (Suzuki *et al.*, 1991), did not show cell adhesion activity, whereas transfectant of cadherin-4, which seems to correspond to chicken R-cadherin, showed strong cell adhesion activity under the same conditions (Tanihara *et al.*, in preparation). Although it seems unlikely, the possibility that the expressed protocadherins were indirectly involved in cell adhesion cannot be ruled out entirely at this time, since the cell adhesion activity was not inhibited with the available antibodies. The cell adhesion sites of classic cadherins are near the N-terminus of the proteins (Nose *et al.*, 1990), and the results of antibody inhibition experiments may suggest that this is also true for the protocadherins. The antibodies we raised are against the EC-3−EC-5 or EC-4−EC-7 of the proteins.

As with the classic cadherins, the cell adhesion activity of protocadherins is likely to be $Ca^{2+}$ dependent. When cell adhesion assay of the protocadherin transfectants was performed using the cell suspensions in HBS with and without $Ca^{2+}$, as described by Urushihara *et al.* (1979), $Ca^{2+}$-dependent cell adhesion was obtained (Sano *et al.*, unpublished observation). The assay was, however, unstable when the cell suspensions were incubated for >1 h. This indicates that cell adhesion assays, at least in some cases, require special procedures. Accordingly, we established a modified cell adhesion assay system. We also have evidence that protocadherins have $Ca^{2+}$ binding activity (Heimark *et al.*, in preparation).

Classic cadherins show homophilic cell adhesion. Considering their unique structural features and cell adhesion properties, the cell adhesion mechanism of protocadherins may be different from that of classic cadherins. As described by Inuzuka *et al.* (1991), i.e. that R-cadherin can interact with N-cadherin, as well as with R-cadherin, there is a possibility that one type of protocadherin may interact with different types of protocadherins, or even the possibility of a ligand/receptor-type interaction. This issue should be studied next.

Recent studies have demonstrated that the cytoplasmic domains of classic cadherins play a pivotal role in cell adhesion activity by virtue of catenin binding and anchorage of cytoskeleton (Gumbiner *et al.*, 1988; Nagafuchi and Takeichi, 1988; Ozawa *et al.*, 1989; McCrea and Gumbiner, 1991; Hirano *et al.*, 1992). Preliminary results showed that these protocadherins did not co-precipitate with catenins in immunoprecipitation.

Even if protocadherins do not serve as the major cell−cell adhesion proteins because of the weak activity, these observations clearly suggest that the protocadherins may be involved directly or indirectly in some cell−cell interaction processes. In this context, it is very suggestive that *Drosophila fat* is involved in the regulation of cell growth, possibly through cell−cell interactions (Mahoney *et al.*, 1991). More studies are apparently necessary to clarify this issue.

Many protocadherins are expressed highly in brain, and each of the neuro-glial cell lines examined so far has expressed at least one protocadherin. Protocadherins are developmentally regulated and their expression is higher in adults than in fetuses or neonates. Moreover, the two protocadherins we discovered were localized to a variety of specific areas in the brain (Davidson *et al.*, in preparation).

Since the central nervous system requires complex cell−cell interactions to develop properly and to maintain its normal structure and function, and since protocadherins seem to be capable of mediating cell−cell interaction, it is tempting to speculate that protocadherins are involved in some highly specialized cell−cell interaction processes in the central nervous system, such as neural network formation.

## Materials and methods

### PCR

PCR was carried out in a manner similar to that described previously (Suzuki *et al.*, 1991). Based on two well-conserved amino acid sequences, $NH_2$-$K^P/_G{}^I/_L D^F/_Y E$ and $NH_2$-NDNAPX, two degenerate oligonucleotides were designed for use as PCR primers: primer 1, 5′-AARSSNNTNGAY-TRYGA-3′ and primer 2, 5′-NNNGGNGCRTTRTCRTT-3′ in which R, S, Y and N designate either A or G, either C or G, either C or T, and either A, C, G or T, respectively. PCR conditions were essentially the same as those described by Saiki *et al.* (1988). Denaturation was performed at 94°C for 1.5 min, annealing at 45°C for 2 min, and polymerization at 72°C for 3 min. The reaction was initiated by adding 2.5 U Taq DNA polymerase, after which 35 reaction cycles were carried out. Taq DNA polymerase and the reaction buffer were obtained from Boehringer Mannheim Biochemicals. After PCR, the reaction products were separated by agarose gel electrophoresis and the DNAs of ~130 bp in size were extracted by a freeze and thaw method. The resultant DNAs were phosphorylated at the 5′ end with T4 polynucleotide kinase and subcloned into the *Sma*I site of M13mp18 by a blunt-end ligation for further analysis.

### Screening of libraries and DNA sequencing

A human brain cDNA library (Stratagene) was screened with $^{32}$P-labeled cDNA probes, human clone #42 and clone #43, obtained by means of PCR using a plaque hybridization method. The positive clones were plaque purified and the inserts were cut out in Bluescript plasmid form using the *in vivo* excision method, then subcloned into the M13 vector for sequencing.

Sequencing of DNAs was carried out according to the dideoxynucleotide chain termination method of Sanger *et al.* (1977), using a sequenase cDNA sequencing kit from United States Biochemicals. DNA and amino acid sequence analyses were carried out using the Beckman Microgenie program.

### Northern analysis

Total RNAs were prepared by the guanidium isothiocyanate method and poly(A)$^+$ RNAs were made by using a FastTrack kit obtained from Invitrogen. RNA preparations were electrophoresed in a 0.8% agarose gel under denaturing conditions, then transferred onto a nitrocellulose filter using a capillary method (Sambrook *et al.*, 1989). Northern blot analyses were performed according to the method of Thomas (1980). The final wash was in 0.2×standard saline citrate containing 0.1% sodium dodecyl sulfate at 65°C for 10 min.

### Plasmid construction and transfection

The full-length cDNAs for pc42 and pc43 were isolated from the respective pBluescript SK(+) clones by digestion with *Ssp*I followed by blunting and digestion with *Xba*I, or double digestion with *Spe*I and *Eco*RV. The pRc/RSV expression vector obtained from Invitrogen was used for transfection experiments. The vectors used for pc42 and pc43 were prepared by digestion with *Hin*dIII, followed by blunt ending and re-digestion with *Xba*I, or by digestion with *Xba*I followed by blunt ending and re-digestion with *Spe*I. The cDNAs were ligated into pRc/RSV vector with T4 DNA ligase, and NM522 cells were transformed with the resultant vectors by a single-step method (Chung *et al.*, 1989). The plasmid DNAs were purified by the CsCl gradient centrifugation method.

L cells were transfected with the resultant constructs by the calcium phosphate method using Pharmacia CellPhect transfection kit and stable transfectants were isolated by G418 selection.

### Preparation of antibodies

The cDNAs corresponding to EC-4−EC-7 of pc42, and the EC3−EC5 of pc43 were prepared by PCR and subcloned into the *Eco*RI−*Xba*I site of an expression vector pMal-cRI (New England Biolabs) in the correct reading frame. The plasmid DNAs were then introduced into NM522 cells by a single-step transformation method. The fusion proteins were induced by the addition of IPTG and were purified from the extract by amylose resin affinity chromatography, as described by the manufacturer (New England Biolabs).

Polyclonal antibodies against pc42 were prepared in rabbits by injecting 0.5 mg of the pc42 fusion protein in Freund's complete adjuvant at four subcutaneous sites. Subsequent injections were in Freund's incomplete adjuvant with 0.5 mg of the fusion protein. The resultant antibodies were purified from the sera with a Sepharose column coupled with the above fusion protein after absorption with a Sepharose column coupled with maltose binding protein (Sambrook *et al.*, 1989).

Monoclonal antibodies against pc43-fusion protein were prepared according to the procedure of Kennett *et al.* (1977). Balb/c mice were immunized with the pc43-fusion protein in Freund's complete adjuvant and subsequent immunizations were in incomplete adjuvant with 100 μg of pc43-fusion protein. Immune spleens were removed and fused with NS1 myeloma cells with PEG 1450. Hybridomas were screened for the production of antibodies in a solid-phase ELISA comparing the difference between wells coated with pc43-fusion protein and maltose binding protein. Hybridoma culture supernatants were further screened for cell surface binding to an astrocytoma cell line UW28 in a solid-phase ELISA and selected for reactivity with astrocytoma membrane proteins by immunoblotting. Positive hybridomas were subcloned twice. The clone 38D10E was used in this study.

### Immunoblotting

Tissue and cell extracts were separated by 7% SDS−PAGE, and the proteins were electrophoretically blotted onto PVDF membranes. The resultant membranes were incubated with 5% skim milk in Tris-buffered saline, pH 8.0 (TBS) for 2 h and then with specific antibodies in the above buffer for 1 h. The membranes were washed three times (5 min each) with TBS containing 0.05% Tween 20 and were incubated with alkaline phosphatase-conjugated anti-rabbit IgG antibody or anti-mouse IgG antibody (Promega) in the same buffer for 1 h. After washing the membranes with TBS containing 0.05% Tween 20, immunolabelled bands were visualized by using Western blue solution obtained from Promega.

### Immunofluorescence microscopy

Cells were grown on a cover slip pre-coated with fibronectin or vitronectin and fixed with 4% paraformaldehyde for 5 min at room temperature or with cold methanol on ice for 10 min followed by 4% paraformaldehyde fixation. After washing with TBS, the cells were incubated with TBS containing 1% bovine serum albumin (BSA) for 30 min and then with the specific antibody in TBS containing 1% BSA for 1 h at room temperature. The cells were then washed with TBS containing 0.01% BSA and incubated with fluorescein isothiocyanate (FITC)-conjugated anti-rabbit antibody or anti-mouse antibody (Cappel) for 60 min at room temperature. The resultant cells were washed again with TBS containing 0.01% BSA and subjected to fluorescence microscopy.

### Cell culture and cell aggregation assay

Cells were cultured in DMEM supplemented with 10% fetal bovine serum at 37°C in a 5% $CO_2$ atmosphere.

Cell aggregation assay was performed according to the method of Urushihara *et al.* (1979), with some modifications. Cells grown near confluence were treated with 0.01% trypsin in the presence of 2 mM $CaCl_2$ or of 1 mM EGTA for 25 min on a rotary shaker at 37°C and collected by centrifugation. The cells were washed three times with $Ca^{2+}$-free Hepes-buffered saline (HBS) after adding soybean trypsin inhibitor, and were resuspended in HBS containing 1% BSA. The cell aggregation was performed by incubating the resultant cells in the 1:1 mixture of DMEM and HBS containing 1% BSA, 2 mM $CaCl_2$ and 20 μg/ml of deoxyribonucelease (Ozawa *et al.*, 1990) on a rotary shaker at 37°C for from 30 min to 3 h.

## Acknowledgements

We thank Drs M.Ozawa and M.Takeichi for their valuable advice and technical help in carrying out this study. We thank Dr E.Ruoslahti for his critical reading and comments on this manuscript. We thank Drs H.Matsumoto, J.P.Thiery, Y.Ohshima and J.M.Kramer for providing us with *Drosophila* cDNA, *Xenopus* cDNA and *C.elegans* cDNA preparations, respectively. We also thank Ms Ann Dawson and Ms Fay Sim for preparing this manuscript. This work was supported in part by grants from NIH (EY08106), the Seaver Foundation and the ICOS Corporation.

## References

Amagai,M., Klaus-Kovtun,V. and Stanley,J.R. (1991) *Cell*, **67**, 869−877.

Bertolotti,R., Rutishauser,U. and Edelman,G.M. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 4831−4835.

Cech,T.R. (1983) *Cell*, **34,** 713–716.
Chen,W.C. and Obrink,B. (1991) *J. Cell. Biol.*, **114,** 319–327.
Chung,C.T.,Niemela,S.L. and Miller,R.H. (1989) *Proc. Natl Acad. Sci. USA*, **86,** 2172–2175.
Detrick,R.J., Dickey,D. and Kintner,C.R. (1990) *Neuron*, **4,** 493–506.
Donalies,M., Cramer,M., Ringwald,M. and Starzinski-Powitz,A. (1991) *Proc. Natl Acad. Sci. USA*, **88,** 8024–8028.
Frixen,U.H., Behrens,J., Sachs,M., Eberle,G., Voss,B., Warda,A., Lochner,D. and Birchmeier,W. (1991) *J. Cell Biol.*, **113,** 173–185.
Fujimori,T., Miyatani,S. and Takeichi,M. (1990) *Development*, **110,** 97–104.
Gallin,W.J., Sorkin,B.C., Edelman,G.M. and Cunningham,B.A. (1987) *Proc. Natl Acad. Sci. USA*, **84,** 2808–2812.
Ginsberg,D., DeSimone,D. and Geiger,B. (1991) *Development*, **111,** 315–325.
Goodwin,L., Hill,J.E., Raynor,K., Raszi,L., Manabe,M. and Cowin,P. (1990) *Biochem. Biophys. Res. Commun.*, **173,** 1224–1230.
Gumbiner,B., Stevenson,B. and Grimaldi,A. (1988) *J. Cell Biol.*, **107,** 1575–1587.
Hatta,K., Nose,A., Nagafuchi,A. and Takeichi,M. (1988) *J. Cell Biol.*, **106,** 873–881.
Hatta,M., Miyatani,S., Copeland,N.G., Gilbert,D.J., Jenkins,N.A. and Takeichi,M. (1991) *Nucleic Acids Res.*, **19,** 4437–4441.
Hirano,S., Kimoto,N., Shimoyama,Y., Hirohashi,S. and Takeichi,M. (1992) *Cell*, **70,** 293–301.
Holton,J.L., Kenny,T.P., Legan,P.K., Collins,J.E., Keen,J.N., Sharma,R. and Garrod,D.R. (1990) *J. Cell Sci.*, **97,** 239–246.
Inuzuka,H., Miyatani,S. and Takeichi,M. (1991) *Neuron*, **7,** 69–79.
Kemler,R, Babinet,C., Eisen,H. and Jacob,F. (1977) *Proc. Natl Acad. Sci. USA*, **75,** 4449–4452.
Kemler,R., Ozawa,M. and Ringwald,M. (1989) *Curr. Opin. Cell Biol.*, **1,** 892–897.
Kennett,R.H., Denis,K.A., Tung,A.S. and Klinman,N.R. (1977) *Curr. Top. Microbiol. Immunol.*, **81,** 77–93.
Kintner,C. (1992) *Cell*, **69,** 225–236.
Koch,P.J., Walsh,M.J., Schmelz,M., Goldschmidt,M.D., Zimbelmann,R. and Franke,W.W. (1990) *Eur. J. Cell Biol.*, **53,** 1–12.
Kozak,M. (1984) *Nucleic Acids Res.*, **12,** 857–872.
McCrea,P.D. and Gumbiner,B.M. (1991) *J. Biol. Chem.*, **266,** 4514–4520.
Mahoney,P.A., Weber,U., Onofrechuk,P. Biessmann,H., Bryant,P.J. and Goodman,C.S. (1991) *Cell*, **67,** 853–868.
Miyatani,S., Shimamura,K., Hatta,M., Nagafuchi,A., Nose,A., Matsunaga,M., Hatta,K. and Takeichi,M. (1989) *Science*, **245,** 631–635.
Nagafuchi,A. and Takeichi,M. (1988) *EMBO J.*, **7,** 3679–3684.
Nagafuchi,A., Shirayoshi,Y.,Okazaki,K., Yasuda,K. and Takeichi,M. (1987) *Nature*, **329,** 341–343.
Napolitano,E.W., Venstrom,K., Wheeler,E.F. and Reichardt,L.F. (1991) *J. Cell Biol.*, **113,** 893–905.
Nose,A., Nagafuchi,A. and Takeichi,M. (1987) *EMBO J.*, **6,** 3655–3661.
Nose,A., Tsuji,K. and Takeichi,M. (1990) *Cell*, **61,** 147–155.
Ozawa,M., Baribault,H. and Kemler,R. (1989) *EMBO J.*, **8,** 1711–1717.
Ozawa,M., Engel,J. and Kemler,R. (1990) *Cell*, **63,** 1033–1038.
Ranscht,B. and Dours-Zimmermann,M.T. (1991) *Neuron*, **7,** 391–402.
Ringwald,M., Schuh,R., Vestweber,D., Eistetter,H., Lottspeich,F., Engel,J., Dölz,R., Jähnig,F., Epplen,J., Mayer,S., Müller,C. and Kemler,R. (1987) *EMBO J.*, **6,** 3647–3653.
Ringwald,M., Baribault,H., Schmidt,C. and Kemler,R. (1991) *Nucleic Acids Res.*, **19,** 6533–6539.
Saiki,R.K., Gelfand,D.H., Stoffel,S., Scharf,S.J., Higuchi,R., Horn,G.T., Mullis,K.B. and Erlich,H. (1988) *Science*, **239,** 487–491.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74,** 5463–5467.
Shimoyama,Y., Yoshida,T., Terada,M., Shimosato,Y., Abe,O. and Hirohashi,S. (1989) *J. Cell Biol.*, **109,** 1787–1794.
Sorkin,B.C., Gallin,W.J., Edelman,G.M. and Cunningham,B.A. (1991) *Proc. Natl Acad. Sci. USA*, **88,** 11545–11549.
Suzuki,S., Sano,K. and Tanihara,H. (1991) *Cell Regul.*, **2,** 261–270.
Takeichi,M. (1991) *Science*, **251,** 1451–1455.
Thomas,P.S. (1980) *Proc. Natl Acad. Sci. USA*, **77,** 5201–5205.
Urushihara,H., Ozaki,H.S. and Takeichi,M. (1979) *Dev. Biol.*, **70,** 206–216.
Vleminckx,K., Vakaet,L.,Jr, Mareel,M., Fiers,W. and Van Roy,F. (1991) *Cell*, **66,** 107–119.
Volk,T. and Geiger,B. (1986) *J. Cell Biol.*, **103,** 1441–1450.

*Received on November 3, 1992; revised on February 25, 1993*