The cadherins comprise a family of single-pass transmembrane proteins critical for cell–cell adhesion in vertebrates and invertebrates. The recently determined structure of the whole ectodomain from C-cadherin suggests that the adhesion of cadherins presented by juxtaposed cells is mediated by a strand-swapped dimer in which core hydrophobic elements are exchanged between the partner molecules. Sequence analysis suggests that several cadherin subfamilies share this adhesive mechanism. Recent work has shed new light on the molecular basis of cadherin adhesion, although understanding the specificity of these interactions remains a major challenge.

Addresses
Department of Biochemistry and Molecular Biophysics, Edward S Harkness Eye Institute, Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA
e-mail: shapiro@convex.hhmi.columbia.edu
Saurabh D Patel and Chien Peter Chen contributed equally to this work

Introduction
The cadherins comprise a family of calcium-dependent cell adhesion molecules that form and maintain adhesive contacts between the cells of solid tissues [1–3]. Many events in the development of complex tissue structures are associated with changes in cadherin expression [4,5,6]. This has led to the prevailing view that selective cell adhesion mediated by cadherins provides a key driving force in the development of tissue architecture [6–8].

Cadherins are single-pass transmembrane proteins characterized by the presence of distinctive cadherin repeat sequences in their extracellular segments [3]. Each of these repeats, consisting of about 110 amino acids, forms a β-sandwich domain with Greek-key folding topology [9–11]. Cadherins typically have several of these ‘cadherin domains’ tandemly repeated in their extracellular segments. The connections between these domains are rigidified by the specific binding of three Ca²⁺ ions between each successive domain pair [12,13]. The presence of Ca²⁺ is necessary for cadherin adhesive function and hence their name arose as an approximate contraction of ‘calcium-dependent adherent protein’.

Cadherins can be classified into several subfamilies [14]: the type I (classical) and type II cadherins, which are ultimately linked to the actin cytoskeleton [15,16]; the desmosomal cadherins (desmocollins and desmogleins), which are linked to intermediate filaments [17,18]; and the protocadherins [19], which are expressed primarily in the nervous system. A large subset of the protocadherins are encoded in novel multigene loci under complex regulation [20,21,22]. In addition, several ‘atypical cadherins’, proteins containing one or more cadherin repeat sequences but bearing no other hallmarks of cadherins, have also been described [14]. A schematic diagram of the domain organization of the major cadherin subfamilies is shown in Figure 1.

Specificity and promiscuity of cadherin–cadherin interactions
The ‘classical’ type I cadherins have been most extensively studied and their binding specificities have been shown to be primarily homophilic [23–25]. That is, expressing a particular cadherin (N-cadherin, for example) will preferentially adhere to other cells expressing the same cadherin, rather than to cells expressing a different cadherin (E-cadherin, for example). However, recent studies indicate that interaction between cadherins can be more promiscuous [26,27,28,29]. Although homophilic binding still appears to be preferred, many type I cadherins can form productive cell adhesive interactions with other type I cadherins. Some type I cadherins (such as N-cadherin and E-cadherin) induce the formation of separate cell aggregates when transfectants are mixed, whereas others (such as N-cadherin and R-cadherin) will co-aggregate. The single in-depth study on the adhesive specificity of type II cadherins yields similar results for this subfamily [27]. Despite the cross-reactivity of cadherins within the same subfamily, it appears that the binding specificities of type I and
type II cadherins are orthogonal. That is to say, type I cadherins may have graded affinities for other type I cadherins and type II cadherins have graded affinities for others of their class [28]. However, current data suggest that type I and type II cadherins do not interact. The ‘orthogonality’ of the type I and type II cadherin adhesion systems may have important consequences for how these molecules function in tissues where many different cadherins may be expressed [5].

In development, cadherins can function in the formation of cell layers (e.g. the separation of the neural tube from the ectoderm, mediated by N-cadherin and E-cadherin) and also in the formation of highly complex tissue structures. Figure 2 shows the expression patterns of four different cadherins in the developing mouse retina, visualized by in situ hybridization [4]. Various stages of development are shown, illustrating the extraordinary temporal and spatial regulation of classical cadherins in...
Type I and type II cadherins each contain five tandem cadherin repeat domains in their extracellular region (EC1–EC5). A key question is to define which of these repeats harbors the cadherin–cadherin binding and specificity sites. This question has been addressed, in one set of experiments, through the study of the behavior of chimeric molecules in cell aggregation assays (Figure 3) [23,26]. In summary, N-cadherin and E-cadherin transfectedants sort to form separate tissue-like cell aggregates, but swapping domains between N-cadherin and E-cadherin can sometimes swap the specificity of cell aggregation. Notably, it is the N-terminal, membrane-distal domain of each molecule (EC1) that carries this property. These experiments provide strong evidence supporting the idea that the adhesive binding site of cadherins is localized primarily within the EC1 domain (see also Update). Additionally, cadherins are synthesized as pro-proteins with a pre-domain at the N terminus of EC1; this must be removed to activate adhesive function [30]. The proximity of this small pre-domain to EC1 provides additional support for the notion that binding and specificity are dominated by interactions involving the EC1 domain and/or regions close to it. Further evidence comes from the work of Engel and collaborators [31,32,33*], who have reported a variety of creative EM studies of cadherin fusion proteins, which have uniformly shown interaction between cadherins at or near their N-terminal domains.

The recent crystal structure of the whole ectodomain (EC1–EC5) from type I C-cadherin [13**] revealed a pair of molecules in a symmetric dimer formed through the interaction of the partner EC1 domains (Figure 4a). The interface is mediated by the exchange of the N-terminal β-strands between the partner EC1 domains. A central feature is the insertion of the conserved Trp2 sidechain from one molecule into the hydrophobic core of the other (Figure 4e). This interface, called a cadherin ‘strand dimer’, had been observed before; structures of EC1 from the classical cadherin N-cadherin revealed an identical configuration [9].

This strand-dimer interface plays a critical role in cadherin-mediated cell adhesion and we are of the opinion that it is the primary adhesive interface of cadherins. Several lines of evidence support this idea. First, W2 and the residues that line its acceptor pocket are highly conserved among cadherins. Second, cadherin-mediated aggregation is abolished (for all cadherins tested thus far) by mutation of this tryptophan to alanine (W2A, Figure 3f,g) or by mutation of one of the alanine residues of the acceptor pocket to encode a larger residue, methionine (A80M, Figure 3h). Such a residue would be expected to block tryptophan insertion [24]. Furthermore, a zebrafish lethal developmental mutation, glass onion, has been shown to encode a W2G mutant of N-cadherin [34*]. Finally, the tryptophan sidechain analog indole-3-acetic acid has been shown to function as an inhibitor of cadherin adhesion, both in cell-based experiments [24] and in experiments with purified protein [35**]. The inability of the related compound 5-methyl indole 3-acetic acid, which modeling studies suggest

**Figure 2**

Illustration of cadherin function in the developing mouse retina. Complex spatial and temporal regulation of cadherin expression is thought to guide the formation of tissue architecture through the differential cell adhesive specificity of different cadherins. The panels show in situ hybridizations for the type I molecules N-cadherin and R-cadherin, and the type II molecules cadherin-8 and cadherin-11, at postnatal days 0, 7 and 21. The classical cadherins are found in the inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; GCL, ganglion cell layer; IPL, inner plexiform layer; INL, inner nuclear layer; OPL, outer plexiform layer; ONL, outer nuclear layer; and OLM, outer limiting membrane. Adapted from [4].

the formation of this complex structure, which consists of multiple cell types arranged in several different cell layers. In light of the large variety of cadherins expressed in such tissues, it seems likely that heterophilic as well as homophilic interactions will be formed.

Despite the importance to development and tissue homeostasis of selective cell adhesion mediated by cadherins, the atomic-level basis of their specificity is still not understood. In this review, we focus on recent high-resolution structural results and discuss their implications for understanding the cell adhesive interactions of cadherins.

**Structural basis of homophilic adhesion**

Type I and type II cadherins each contain five tandem cadherin repeat domains in their extracellular region (EC1–EC5). A key question is to define which of these repeats harbors the cadherin–cadherin binding and specificity sites. This question has been addressed, in one set

The recent crystal structure of the whole ectodomain (EC1–EC5) from type I C-cadherin [13**] revealed a pair of molecules in a symmetric dimer formed through the interaction of the partner EC1 domains (Figure 4a). The interface is mediated by the exchange of the N-terminal β-strands between the partner EC1 domains. A central feature is the insertion of the conserved Trp2 sidechain from one molecule into the hydrophobic core of the other (Figure 4e). This interface, called a cadherin ‘strand dimer’, had been observed before; structures of EC1 from the classical cadherin N-cadherin revealed an identical configuration [9].

This strand-dimer interface plays a critical role in cadherin-mediated cell adhesion and we are of the opinion that it is the primary adhesive interface of cadherins. Several lines of evidence support this idea. First, W2 and the residues that line its acceptor pocket are highly conserved among cadherins. Second, cadherin-mediated aggregation is abolished (for all cadherins tested thus far) by mutation of this tryptophan to alanine (W2A, Figure 3f,g) or by mutation of one of the alanine residues of the acceptor pocket to encode a larger residue, methionine (A80M, Figure 3h). Such a residue would be expected to block tryptophan insertion [24]. Furthermore, a zebrafish lethal developmental mutation, glass onion, has been shown to encode a W2G mutant of N-cadherin [34*]. Finally, the tryptophan sidechain analog indole-3-acetic acid has been shown to function as an inhibitor of cadherin adhesion, both in cell-based experiments [24] and in experiments with purified protein [35**]. The inability of the related compound 5-methyl indole 3-acetic acid, which modeling studies suggest
could not be accommodated in the W2 acceptor pocket, to inhibit adhesion provides evidence for the specificity of this inhibition [24].

Cadherins may be able to dimerize in cis orientations (between molecules presented on the same cell), as well as trans (molecules from different cells). Interestingly, biochemical studies employing precipitation of epitope-tagged cadherins from transfected cells indicate that both cis and trans dimers do indeed form, and that they arise through the use of the same interface residues [26,36,37]. Although this may seem, at first, counterintuitive, the long somewhat flexible nature of the five-domain cadherin extracellular region appears to put little constraint on this alternative usage, which may play an important role in the biology of cadherins. Specifically, it is possible that cadherins may exist as cis strand dimers on a single cell surface, and that these cis dimers may be replaced by or exist in equilibrium with trans strand dimers in the presence of a juxtaposed cadherin-presenting cell.

Cadherins mediating cell adhesion must bind together from two opposing cells. Thus, the adhesive bond must form at the cell surface between cadherins that have been synthesized and processed in different cells. It is not then surprising that monomeric forms of cadherins should exist and indeed have been observed in high-resolution structural studies [10,12,24,32]. These monomeric forms reveal properties of the adhesive interface that are informative about potential mechanisms of adhesion. Two monomer forms of EC1 have been observed: a structure with the A-strand (which encompasses W2) disordered and the W2 acceptor pocket exposed to solvent (Figure 4b,c); and the A-strand closed form (Figure 4d), in which W2 inserts into its own acceptor pocket rather than that of a partner protomer. This plasticity of the A-strand suggests unique properties that are likely to be important for strand-dimer formation between cadherins presented by juxtaposed cells. Two crystal structures and an E-cadherin EC1 NMR structure reveal that the A-strand, which is required to complete the β-sandwich fold of EC1, is disordered. As a result, conserved hydrophobic residues are exposed to solvent. This conformational heterogeneity suggests intrinsic instability of the A-strand in the monomer, as it is disordered in some structures and bound in others. The absence of a stable structure should

Figure 3

Transfected cell experiments can be used to study the adhesive properties of cadherins. L-cells (mouse fibroblasts) become adhesive when stably transfected with classical cadherins (a and e). Transfectants can be labeled with lipophilic dyes so that they can be distinguished after mixing. N-cadherin transfectants adhere to N-cadherin transfectants (a), but sort out from E-cadherin transfectants (b). This adhesive specificity appears to reside in the EC1 domain; a chimeric molecule in which the EC1 domain of N-cadherin replaces the EC1 domain of E-cadherin acquires the adhesive specificity of N-cadherin (c) and effectively sorts out from E-cadherin (d). The strand-dimer interface is essential for cadherin cell adhesive function. Mutation of either W2 (f,g) or residues in its acceptor pocket (h) abolishes adhesive function and also abolishes the ability of these proteins to interact with wild-type cadherins (h). Adapted from [24,26].
favor the ejection of the critical W2-containing A-strand from the protomer body and is likely to be a key element driving the formation of strand-dimer interactions. It should be noted, as a caveat, that some of the monomeric cadherin structures were obtained from cadherin fragments that contained extra amino acids at the N terminus as the result of cloning artifacts. This could be significant, as cell-based studies have shown that inexact cleavage of the pre-domain (leaving as few as four amino acids extra at the N terminus) can abolish the adhesive function of E-cadherin [30]. The kinetics of association of individual E-cadherin fragments have been studied by flow chamber analysis [35**]. Association of these purified proteins was shown to be dependent on W2 and inhibited by indole 3-acetic acid. Furthermore, the bond duration was remarkably brief, estimated at approximately 2 s. Although the full implications of this rapid exchange are not yet clear for cadherin function at the cellular level, this observation may provide a rationale for the apparent need for the clustering of cadherins at cell-cell contacts [38,39].

Although we think that the majority of data indicate that EC1 is the primary site of adhesive binding in cadherins, surface force experiments [40,41] and some cell-based assays [42] suggest the involvement of other domains. In the surface force studies, the extracellular portions of C-cadherin molecules are bound to two opposing plates through C-terminal 6-His tags. These plates are then brought together and the force required to separate them is measured as a function of the interplate distance. The results of these experiments revealed force maxima at several discrete distances. This led the investigators to suggest that different cadherin repeats contribute directly to adhesion through multiple binding sites that form distinct interdigitated conformations to produce complexes of different lengths. In a similar vein, cell aggregation

---

**Figure 4**

High-resolution structures of cadherin extracellular regions. (a) Crystallographically observed dimer of the ectodomain from C-cadherin, which comprises five tandem cadherin repeats, drawn as a Cα worm. The dimer interface joins EC1 domains from each molecule. The cadherin ectodomains are oriented as if emanating from opposing cell surfaces. Note that three Ca²⁺ ions (green spheres) are ligated at the interface between successive domains, rigidifying the structure of each ectodomain, which is highly curved. Two disulfide bonds, drawn in magenta, are found at the base of EC5, near the transmembrane attachment point. Some cadherin structures have revealed alternative monomeric conformations of EC1. GRASP surfaces of EC1 are shown for structures determined from two-domain cadherin fragments. (b) N-cadherin, in which the N-terminal A-strand, including W2, is disordered and (c) E-cadherin, in which the A-strand is also disordered. In these structures, the hydrophobic pockets are exposed to solvent (red arrows). (d) An E-cadherin structure, in which the W2 sidechain is inserted into the hydrophobic pocket of its own protomer. (e) Detailed stereo view of the strand-dimer interface. This is a twofold symmetric interaction, in which the W2 sidechain from one EC1 domain (green) inserts into the hydrophobic pocket presented by its partner EC1 domain (gray). A conserved salt bridge is formed between the N-terminal amino group and the sidechain of Glu89. The hydrophobic pocket is lined by conserved residues Tyr36 and Ile24, and positions 80 and 78 are occupied by residues with small sidechains in order to accommodate the tryptophan insertion. Panels (a and e) were adapted from [13**].
One possibility is that additional domains are required for assays is not evident from the crystallographic studies. For N-cadherin, the EC1–EC2 truncated form can mediate strong adhesion (DR Colman, W Shan, personal communication). The molecular basis of the requirement for at least two domains in cell aggregation essays is not evident from the crystallographic studies. One possibility is that additional domains are required for a minimum clearance from the cell surface, which of course contains myriad other proteins. Another possibility involves a potential active role for Ca$^{2+}$ ligation (which requires more than the EC1 domain) in inducing a conformational change to an adhesion-competent state. It is of interest in this regard that an NMR study revealed chemical shift changes in residues lining the W2 acceptor pocket upon Ca$^{2+}$ ligation, even though these residues are at least 20 Å away from the Ca$^{2+}$-binding sites [10,43]. This explanation does not, however, explain the surface force experiments. A third possibility is that other molecular interfaces, yet to be identified, are involved in adhesion. The primary difficulty with this idea comes from the domain-swap experiments, which indicate that binding specificity is localized exclusively to EC1.

**Potential specificity determinants**

Multiple sequence alignment of EC1 domains from representative cadherins (Figure 5) illustrates the highly conserved nature of W2 and residues lining its hydrophobic acceptor pocket. The protocadherins do not share this conservation, suggesting the possibility that they may use a different interface to mediate adhesion. Nonetheless, the strand dimer probably represents a binding mode that is common to type I, type II and desmosomal cadherins. Although peripheral residues may change

---

**Figure 5**

![Multiple sequence alignment and pairwise sequence identities of EC1 domains.](image)

Multiple sequence alignment and pairwise sequence identities of EC1 regions of representative cadherin superfamily members. (a) Highlighted in green are the conserved Ca$^{2+}$-binding sites, which provide the most recognizable sequence determinants of cadherin sequence repeats. Red highlighted positions indicate conserved tryptophan residues: W2 constitutes a central element of the adhesive interface in type I cadherins and W4, which is conserved in type II cadherins, may be involved in adhesion in this cadherin subclass. Residues in magenta line the hydrophobic pocket that accommodates the conserved W2 in the classical cadherins. Highlighted in gray are analogous hydrophobic pocket positions for type II and desmosomal cadherins, inferred from sequence analysis. The blue highlighted position shows a conserved glutamic acid residue that is involved in forming a salt bridge with the N terminus in the adhesive interface of type I cadherins. Above the alignment are the β-strand positions determined by crystallographic analysis of C-cadherin. (b) Within each cadherin subfamily, there is high sequence identity. Pairwise sequence alignments were obtained using CLUSTALW 1.75 with default gap penalty parameters and the BLOSUM series of scoring matrices. Pcdh, protocadherin.
during evolution to provide binding specificity, the major elements of the cadherin–cadherin interface remain unchanged. Thus, cadherins appear to be different from other classes of adhesion molecules, such as those of the immunoglobulin superfamily [44], which display many structurally disparate modes of interaction.

The conservation of the central elements of the adhesive interface provides a simple explanation for the cross-reactivity of binding among subfamily members. On the other hand, there must be substantial differences between the binding interfaces of type I and type II cadherins, as they do not bind to one another. Indeed, type II cadherins show distinct differences from canonical type I sequences. Most notably, as is evident in Figure 5a, type II cadherins have two conserved tryptophan residues in their A-strands (W2 and W4), rather than a single conserved tryptophan. Additional potentially significant differences between type I and type II cadherins are suggested by crystal structures of type I cadherins. For example, the sidechain at position 23 of type I cadherins is partially buried upon strand-dimer formation and is involved in an intermolecular hydrogen bond with the sidechain at position 8; these residues are always polar in type I cadherins. However, most type II cadherins have phenylalanine at position 8 and a positively charged residue at position 23, so that a comparable hydrogen-bonding interaction is not possible. Understanding the significance of these differences must await the determination of crystal structures of type II cadherins.

Possible specificity determinants for cadherins within a single subfamily (i.e. type I to type I interactions) include those residues present in the adhesive interface, excluding the conserved W2 and small hydrophobic pocket residues, which are common to all subfamily members. Currently, structures of dimeric conformations are available only for C-cadherins and N-cadherins [9,13**]. With this limited set of structures, it is difficult to extract principles that could explain specificity among type I cadherins generally, which show graded levels of affinity for one another.

Conclusions
Cadherins appear to function in cell adhesion through a common mechanism centered around β-strand exchange between monomers and the insertion of hydrophobic residues (W2 in the case of type I cadherins) into the core of the partner molecule. This mechanism of interaction appears to be common to cadherins known to function in cell adhesion and specificity has probably arisen through changes in residues peripheral to these central interface elements. Cadherin EC1 domains have unique properties that destabilize ‘closed’ conformations and promote the partial disassembly of the domain (by ejection of the A-strand), thus enabling strand-dimer formation. The human genome contains about a hundred cadherin genes and it seems likely that those that function in intercellular adhesion do so through this common mechanism. Evidence for this can be seen in the amino acid sequences of type I, type II and desmosomal cadherins, but not the protocadherins. Thus, structural analysis of protocadherins is a high priority in order to gain a better understanding of the cadherin adhesive repertoire.

Our current understanding of cadherin specificity is poor at best. Although we can now be reasonably confident about the identity of the interface that mediates cadherin adhesion, little is known about the factors determining the binding preferences of cadherins. Additional structural studies will undoubtedly shed light on this issue, but these will need to be augmented by a variety of computational tools. There has been considerable progress in our understanding of the energetic basis of protein–protein interactions [45] and it will be of interest to exploit what has been learned to elucidate the design principles used by cadherins to achieve specificity. Some of the questions that need to be addressed for cadherins are quite general, as it is not uncommon for protein–protein recognition modules to share common affinity-determining elements while achieving specificity in other regions of the interface.

Update
Using electron tomographic reconstructions of plastic sections of neo-natal mouse skin, He et al. [48**] visualized the organization of desmosomes in situ at resolutions corresponding to between ~12 and ~30 Å. The crystal structure of the ectodomain from C-cadherin could be modeled into the tomographic density with good fit. This study clearly shows that the predominant site of interaction between the desmosomal cadherin molecules is the EC1 domain. Furthermore, this study presents evidence for the presence of both cis and trans dimers mediated through the EC1 domain, consistent with previous structural and biochemical analyses of the strand dimer.

References and recommended reading
Papers of particular interest, published within the annual period of review, have been highlighted as:
• of special interest
•• of outstanding interest

This paper compares the expression patterns of type II cadherins in the developing spinal cord of chick embryos. The authors demonstrate, for the first time, extracellular markers of different motor neuron pools, which have previously only been delineated by the expression patterns of transcription factors. They also show that cadherins can drive the segregation of neurons into different motor pools by the ectopic expression of MN-cadherin in the chick embryo.


The first high-resolution crystal structure of the whole functional ectodomain of a type I cadherin, C-cadherin, is presented. The structure shows an adhesive interface involving the N-terminal cadherin domains of two cadherin protomers. This structure supports the strand-dimer model of the binding interaction between cadherin molecules.


Before the publication of this paper, there had been several conflicting reports as to whether or not the protocadherin gene cluster was regulated by trans splicing. These authors show, through a number of beautiful experiments, that intracodon splicing occurs via a conventional cis splicing mechanism.


The authors performed adhesion flow assays using purified cadherin proteins and cells expressing cadherins. They observed that none of the cadherin-expressing cells exhibited substantial adhesive specificity for either of the two purified cadherin proteins tested in the assay. This study indicates that classical cadherins may be far more promiscuous in their binding interactions with other classical cadherins than was previously thought.


This EM study of the complete ectodomain of a type II cadherin fused to the oligomerization domain of cartilage matrix protein reveals Ca 2+ dependent ring-like and double ring-like arrangements involving the EC1 cadherin domains. Similar ring-like structures were observed in earlier analogous EM studies of type I cadherins. This suggests a common mechanism of adhesion for type I and type II cadherins.


The authors use genetic screens to show that the glio117 allele encodes a W2G substitution in the N-cadherin gene. This mutant allele, which is embryonically lethal, causes severe disruption of neural architecture. This paper indicates the critical role of the W2 amino acid in the EC1 domain of N-cadherin.


This is the first reported study to provide quantitative data on the timescale of the kinetics of the binding interaction between cadherins, using an in vitro flow chamber. Glass beads coated with cadherin EC1-EC2...
fragments were sent through a laminar flow chamber also coated with cadherin EC1-EC2 fragments. A camera was used to monitor the progress of a glass bead; binding timescales were calculated based upon how long a bead was retarded in the chamber. The unstressed lifetime of individual E-cadherin interactions was as brief as 2 s.

36. Klingelhofer J, Laur OY, Troyanovsky RB, Troyanovsky SM: Dynamic interplay between adhesive and lateral E-cadherin dimers. Mol Cell Biol 2002, 22:7449-7458. This study employed a mixed culture co-immunoprecipitation assay involving cells expressing E-cadherin tagged with either Myc or Flag epitopes. Ca2+ -binding site mutants of E-cadherin were constructed, and cells expressing these mutants can form both adhesive (trans) dimers and lateral (cis) dimers. A W2A mutant formed neither type of dimer. This suggests that the interface mediating cis and trans dimers may be one and the same.


