Structure, Vol. 12, 793–805, May, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.str.2004.02.034

Structure of the Neural (N-) Cadherin Prodomain Reveals a Cadherin Extracellular Domain-like Fold without Adhesive Characteristics

Alexander W. Koch,[1,3,4,*] Amjad Farooq,[2] Weisong Shan,[1,3] Lei Zeng,[2] David R. Colman,[1,3] and Ming-Ming Zhou[2,*]
[1]Fishberg Research Center for Neurobiology
[2]Structural Biology Program
Department of Physiology and Biophysics
Mount Sinai School of Medicine
One Gustave L. Levy Place
New York, New York 10029
[3]Montreal Neurological Institute
McGill University
3801 University Street
Montreal, Quebec H3A 2B4
Canada

Summary

Classical cadherins mediate cell-cell adhesion through calcium-dependent homophilic interactions and are activated through cleavage of a prosequence in the late Golgi. We present here the first three-dimensional structure of a classical cadherin prosequence, solved by NMR. The prototypic prosequence of N-cadherin consists of an Ig-like domain and an unstructured C-terminal region. The folded part of the prosequence—termed prodomain—has a striking structural resemblance to cadherin "adhesive" domains that could not have been predicted from the amino acid sequence due to low sequence similarities. Our detailed structural and evolutionary analysis revealed that prodomains are distant relatives of cadherin "adhesive" domains but lack all the features known to be important for cadherin-cadherin interactions. The presence of an additional "nonadhesive" domain seems to make it impossible to engage homophilic interactions between cadherins that are necessary to activate adhesion, thus explaining the inactive state of prodomain-bearing cadherins.

Introduction

In metazoans, cadherins are Ca^{2+}-dependent mediators of intercellular adhesion. They operate in a wide variety of physiological settings, comprising embryogenesis and organogenesis (Gumbiner, 1996; Takeichi, 1995; Tepass et al., 2000). They also play a dominant role in pathological conditions where cell adhesion is impaired. For instance, loss of, or a switch in, cadherin subtype expression has been attributed to tumor malignancy (Cavallaro et al., 2002). Certain pathogens use cadherins to enter host cells (Mengaud et al., 1996; Schubert et al., 2002) or to spread between cells (Sansonetti et al., 1994), and defective cadherins are the cause of several severe skin diseases (Lin et al., 1997). The widespread biological role of cadherins can be attributed to several important characteristics of these molecules. First, cell adhesion mediated by cadherins is generally homotypic; that is to say that cells expressing a given cadherin only adhere to cells expressing the same cadherin, thereby providing adhesive specificity (Takeichi, 1988). Although heterophilic binding modes between cadherin molecules have been reported (Shan et al., 2000; Shimoyama et al., 2000; Volk et al., 1987), the stronger homophilic binding seems to be the preferred interaction (Kemler, 1992; Nose et al., 1990). Secondly, cadherins have distinct cellular expression patterns that are perhaps best developed in the vertebrate nervous system, where cadherin expression sharply reflects the regional and functional differentiation of the CNS (Redies, 2000). Finally, cadherins exist as a superfamily whose diverse functions are not only a consequence of differences within their extracellular domains, but may be attributed to the variety of cytosolic binding partners that engage cadherins in signaling events (Aberle et al., 1996; Yap et al., 1997a).

Based on functional differences, sequence similarities, and the number of extracellular cadherin domains (CADs), the cadherin superfamily can be subdivided into six subfamilies (Nollet et al., 2000): classical (type I) cadherins, type II cadherins, desmocollins, desmogleins, protocadherins, and cadherin-related proteins. Classical cadherins (e.g., N-, E-, and C-cadherin) consist of five CADs, a single transmembrane helix, and a highly conserved cytosolic domain. Structures have been solved for single, N-terminal CADs of E-cadherin (Overduin et al., 1995) and N-cadherin (Shapiro et al., 1995); for N-terminal domain pairs of the same cadherins (Nagar et al., 1996; Pertz et al., 1999; Tamura et al., 1998); for the cytosolic domain of E-cadherin in complex with its binding partner β-catenin (Huber and Weis, 2001); and recently for the entire extracellular region (ectodomain) of C-cadherin (Boggon et al., 2002). These studies greatly enhanced our understanding of cadherin-based adhesion, revealing for instance that cadherins must undergo both *cis*- and *trans*-interactions to form functional adhesive contacts; however, none of the existing structural data completely explains the "adhesive" behavior of cadherins.

Cadherin-based adhesiveness is posttranslationally regulated through signaling events that involve cadherin cytoplasmic domains and their binding partners, e.g., catenins (Aberle et al., 1996; Gumbiner, 2000; Pokutta and Weis, 2002). In addition, the formation of *cis*-interactions is believed to be critical for regulating the adhesive strength of cadherins (Brieher et al., 1996; Pertz et al., 1999; Shan et al., 2000; Takeda et al., 1999; Tanaka et al., 2000; Yap et al., 1997b). Another way by which cadherin function is regulated is through proteolytic processing (Ozawa and Kemler, 1990), as many cadherins have prosequences N-terminal to "adhesive" domains. These sequences are cleaved off by an endoprotease, probably within the late Golgi and after catenins have

*Correspondence: akoch@gene.com (A.W.K.), ming-ming.zhou@mssm.edu (M.-M.Z.)
[4]Present address: Department of Protein Chemistry, Genentech Inc., South San Francisco, California 94080.

already been attached, thus activating cadherins for adhesion (Ozawa, 2002; Ozawa and Kemler, 1990; Shore and Nelson, 1991; Wahl et al., 2003). Ozawa and Kemler demonstrated that correct processing of E-cadherin is neither required for biosynthesis nor for transport of the protein to the cell surface but rather important for its adhesive function (Ozawa and Kemler, 1990). These observations, and the fact that some but not all cadherins have large prosequences close to homophilic interactions sites, spurred us to perform a detailed structural analysis of the prosequence of N-cadherin.

The high-resolution three-dimensional NMR structure of the N-cadherin prosequence (NPro), which we present here, shows that the core region of NPro is a folded domain and, remarkably, has a cadherin-like fold, unpredictable from the primary sequence. Nevertheless, NPro lacks all the structural features known to be required for cadherin "adhesive" interactions, including an absolutely conserved Trp that is engaged in homophilic interactions and all the calcium binding sites. This, together with the close physical proximity of NPro to cadherin "adhesive" domains in situ, explains why L cells expressing uncleaved NPro do not coaggregate, despite the surface localization of prodomain-containing N-cadherin. The present work provides first structural insights for the maintenance of the "nonadhesive" state of classical cadherins through retention of their prodomains.

Results and Discussion

Prosequences within the Cadherin Superfamily

A sequence analysis of protagonists from all the major cadherin subfamilies shows that the subdivision of the cadherin superfamily is well reflected by the presence and size of prosequences (Figure 1). In addition, proteolytic cleavage sites are well conserved within each individual subfamily. All cleavage sites contain a dibasic recognition motif that is generally recognized by members of the furin protease family (known as endoprotein convertases), which process a wide range of extracellular proteins, growth factors, and hormones (Seidah and Chretien, 1999). It has been shown in cell transfection assays that furin can process E-cadherin, but other proteases can also cleave, as E-cadherin is properly processed even in a furin-deficient cell line (Posthaus et al., 1998). A cadherin-specific convertase, therefore, has yet to be identified.

Notably, only cadherins with five extracellular cadherin domains (CADs) possess prosequences, and among them only classical (type I) cadherins (including T-cadherin) and desmocollins have prosequences of considerable length. Prosequences from these cadherins share significant sequence similarities, especially within a 90 amino acid core region, ranging from Cys32, which is absolutely conserved, to Leu121 (Figure 1B). Cadherin prosequences, however, have no significant similarities to other sequences of the database. Therefore, we chose the prosequence of N-cadherin (NPro) as a prototype of a family of sequences for further structural and functional analyses.

Proteolytic Processing Activates Cadherins

Ozawa and Kemler previously demonstrated in an elegant way—by mutating the furin-type cleavage site between the prosequence and first cadherin domain—that E-cadherin with an uncleaved prosequence is efficiently expressed on cell surfaces, forms complexes with catenins, and shows the typical Ca^{2+}-dependent resistance to proteases (Ozawa and Kemler, 1990). E-cadherin's adhesive activity, on the other hand, was completely abolished. Using these prior experiments as a guide, we examined the localization of prosequence-bearing N-cadherin in transfected L cells, using a specific NPro antibody that we generated in this study.

Immunostaining of L cells, transfected with the full-length N-cadherin cDNA (L_N cells), with affinity-purified NPro antibodies and a specific N-cadherin antibody illustrates that N-cadherin is expressed on the cell surface, especially at cell-to-cell contact sites, whereas NPro is localized inside the cells, especially within Golgi-like structures (Figure 2B). When L cells were transfected with an N-cadherin mutant (L_{ProN} cells), in which the endogenous cleavage site was mutated into a factor Xa cleavage site (Figure 2A), prosequence-containing N-cadherin was surface expressed as much as the wild-type (Figure 2B). In cell aggregation assays, L_N cells typically form large aggregates (Figure 2C, first column), whereas L_{ProN} cells form only a few, very small aggregates similar to untransfected, "nonadhesive" L cells. After removal of the prosequence by factor Xa, L_{ProN} cells showed similar aggregation tendencies as L_N cells (Figure 2C, third column). Removal of the prosequence by trypsin had the same effect, demonstrating that other potential trypsin cleavage sites are well protected and, therefore, the cadherins are properly folded.

These results clearly demonstrate that the presence of NPro at the cell surface confers "nonadhesivity," but does not affect the processing and surface expression of wild-type N-cadherin. The fact that the adhesive function of L_N cells can be readily rescued through subsequent removal of NPro via exogenous factor Xa supports the notion that the prosequence does not influence folding or posttranslational modification of N-cadherin. These results agree with the earlier data obtained for E-cadherin (Ozawa and Kemler, 1990), but additionally, we were able to visualize directly the prosequence with specific NPro antibodies.

Structure Determination of the N-Cadherin Prodomain

The recombinant protein construct we designed for NMR structural analysis comprises the entire prosequence of N-cadherin without the signal peptide (residues Glu24–Arg159). The construct also contains an N-terminal histidine tag to facilitate purification by affinity chromatography. Protease resistance, CD spectra, and ^{15}N-HSQC experiments (data not shown) all revealed that the N-cadherin prosequence contains a well-folded three-dimensional structure. This and the subsequent solution structure (see below) demonstrate for the first time that the N-cadherin prosequence, and most likely all classical cadherin prosequences (Figure 1B), constitutes an independently folded protein domain rather than a nonstructured peptide. Therefore, we denote here the N-cadherin prosequence as N-cadherin prodomain (NPro).

A

Cadherin subfamily	Pro-sequences	Consensus cleavage site
Classical (type I) cadherins e.g. N-, E-, C-, R-, P-cadherin	~130 aa	L-X_1-R-X_1-K-R↓X_2-W (X_1= R, Q, K; X_2=D, E)
Type II cadherins e.g. VE-cadherin, cadherin-6, 8, 10-12	20-35 aa	R-X_1-K-R↓X_2-W (X_1= R, Q, K; X_2= G, S, D)
Desmocollins e.g. Dsc 1-3	~105 aa	R-X_1-K-R↓X_2-W (X_1= A, S, T; X_2= R)
Desmogleins e.g. Dsg 1-3	~ 25 aa	R-X_1-K-R↓X_2-W (X_1= Q, T; X_2= E, A)
Protocadherins e.g. α-, β-, γ-subclasses, protocadherin-8	—	—
Cadherin-related proteins e.g. DE-, DN-cadherin, Ret, Fat, HMR1, Flamingo cadherins	—	—

B

```
                30        40        50        60        70        80        90
N-CAD:  ..EASGEIALCKTGFPEDVYSAVLPKDVHE.GQPLLNVKFSNCNRKRKVQYESSEPADFKVDEDGTVYAVRSFP
R-CAD:  HREDLTVREACKAGFSEEGYTALISPNVLE.GEKLLKVEFSSCVGTKGMQYETNS.LDFKVGADGTVFATRELK
C-CAD:  .....DV.SGCKPGFSSAEYIFSVNRRELERGRKLGKVNFSDCTTRKHGLYDVGD.SRFRVLPDGTVLVKRHVK
E-CAD:  ...ELEP.ESCSPGFSSEVYTFPVPERHLERGHVLGRVRFEGCTGRPRTAFFSED.SRFKVATDGTITVKRHLK
T-CAD:  ......DDLECTPGFQRKVLHIHQPAEFIE.DQPVLNLTFNDCKGNEKLHYEVSS.PHFKVNSDGTLVALR..N
P-CAD:  ........EPCRAVFREAEVTLEAGGAEQEPGQALGKV.FMGCPGQEPALFSTDN.DDFTVRNGETVQERRSLK
DSC3:   ..........CKK.......VTFHVPSTLEVDKIIGRVSLKECLSSADGIMPSDP..DFRVLDDGSVYPTRAVV
DSC1:   .......CDACQK.......VSLHVPSHLKAETPVGKVNLEECLKSPSLILSSDP..AFRILEDGTIYTTHDLL
DSC2:   ......ASDACKE.......VVLQVPSELPAEKFVGRVNLMDCLKSADIVHLSDP..DFQVLEDGSVYTTSSVV

             100       110       120       130                 140       150
N-CAD:  LTAEQAKFLIYAQDKETQEKWQVAVN.LSREPTLTEEPMKE........PHEIEEIVFPRQLAKHSGALQRQKR
R-CAD:  IPSEQVAFTVTARDRQSAEQWAAMVRLLVAQTSSAHSEHKKGQTVALDPSQPPNDTILPWPQHQSSGGLRRQKR
C-CAD:  LH.KDTKFTISTWDARGIKHSTNIA..VASK...RHRSGEE......AHSRSSKLPVLTFPETHT..GLKRKKR
E-CAD:  LHKLETSFLVRARDSSHRELSTKVT..LKSMGHHHHRHHHR......DPASESNPELLMFPSVYP..GLRRQKR
T-CAD:  ITAVGRTLFVHARTPHAEDMAELVIV.....................GGKDIQGSLQDIFKFARTSPVPRQKR
P-CAD:  ER...NPLKIFPSKR.....................................................ILRRHKR
DSC3:   LSDEKRSFTIQLSDSKMQTQKEIPVI.....................LEHK.KKVLKKRHTKETVLRRSKR
DSC1:   LSSEKRGFSILLSDGQGQEQKKLEVV.....................LSAREKKVFRKRHTKEPVHNRSKR
DSC2:   LSSGQRSFTIWLFSTDSQEEREISVH.....................LEGP.VEVLNKRPHTEKVLSRAKR
```

Figure 1. Comparison of Cadherin Prosequences

(A) Occurrence, length, and consensus protease recognition sites of prosequences within the cadherin superfamily. Lengths of prosequences are shown by the number of amino acids (aa), and protease cleavage sites are indicated by arrows.

(B) Sequence alignment of cadherin prosequences. Sequences of six classical cadherins and three desmocollins were aligned, and the positions of secondary structure elements from the NPro structure (see below) are indicated by yellow shading. Conserved identical residues are shown in red, and conserved similar residues in orange. Single-letter code for amino acids was used, and abbreviations for cadherins are as follows: N-, neural; E-, epithelial; C-, (*Xenopus*) compaction; R-, retinal; B-, (chicken) blastomer; VE-, vascular endothelium; Dsc, desmocollin; Dsg, desmoglein; DE-, *Drosophila* epithelial; DN-, *Drosophila* neuronal; Ret, from *ret* protooncogene; Fat, from *fat* tumor suppressor gene; HMR1, from the *C. elegans* humpback gene (*hmp*) related gene *hmr-1*.

We determined the structure of NPro from a total of 2129 NMR-derived distance, torsion angle, and hydrogen bonding restraints (Table 1) using standard heteronuclear multidimensional NMR methods (Clore and Gronenborn, 1994; Sattler et al., 1999). While the majority of NOE (nuclear Overhauser effect) peaks were manually assigned to obtain distance restraints (Table 1), the program ARIA was used to assign additional NOE peaks in the structure refinement. Figure 3A depicts the superposition of backbone atoms (residues 31–124) of the 20 lowest energy structures derived from NMR data. The C-terminal part of NPro (residues 125–159), which consists of a low homology region followed by a proteolytic cleavage site (Figure 1B), was unstructured since it lacks discernable long-range NOEs and showed very narrow chemical shift dispersion in the NMR spectra. Within the

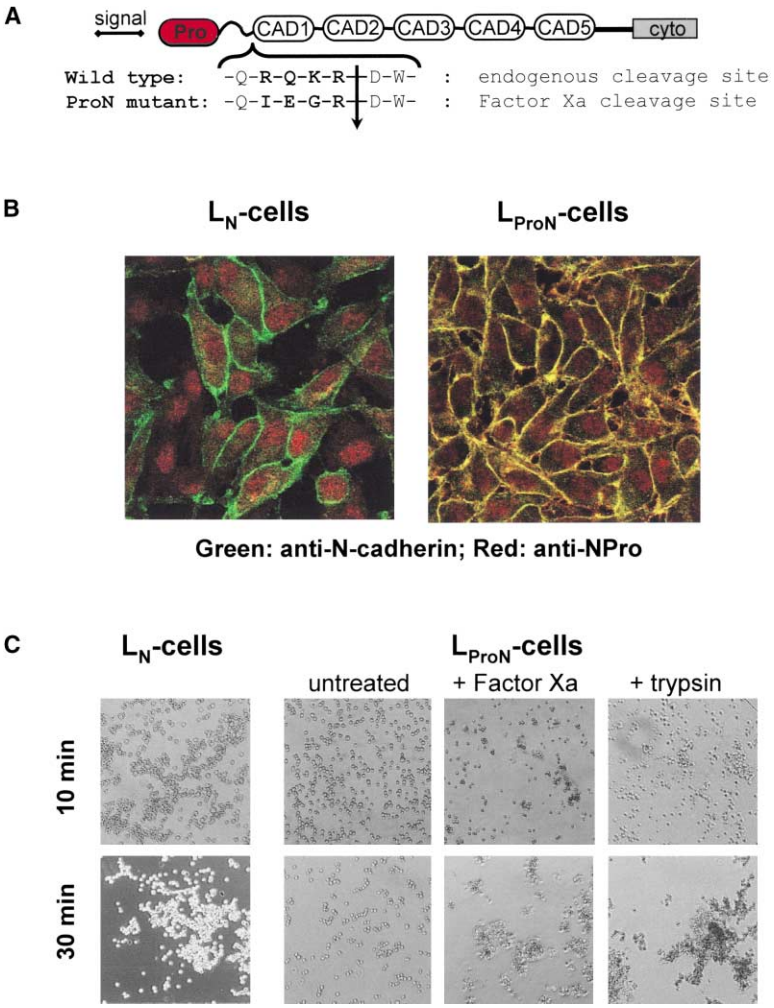

Figure 2. Immunostaining and Cell Aggregation Assays of L_N Cells and L_{ProN} Cells

(A) Domain organization of N-cadherin with the wild-type cleavage site and the cleavage site of the ProN mutant. The prodomain is shown in red, cadherin domains are depicted as white ovals, and the cytosolic domain is shown as a gray rectangle. In the ProN mutant, the wild-type cleavage site was changed to a factor Xa cleavage site as indicated.

(B) Immunostaining of L cells transfected with wild-type N-cadherin (L_N cells) or with the ProN mutant (L_{ProN} cells). N-cadherin staining with rat anti-NCAD2 antibodies is shown in green, and prodomain staining with rabbit anti-NPro antibodies is shown in red.

(C) Cell aggregation assays with L_N and L_{ProN} cells. Cells were dissociated and allowed to reaggregate in the presence of Ca^{2+} over time as indicated.

sequence range of residues 31–124, all the structures are well defined, and the root mean square (rms) distributions for backbone and all heavy atoms for the structured region are 0.67 ± 0.08 Å and 1.29 ± 0.08 Å, respectively (Table 1). Interestingly, we found that the protein contains a disulfide bond formed between Cys32 and Cys63, which was established through the characteristic chemical shift value for the β-carbon of Cys32 and confirmed by mass spectrometry analysis of NPro fragments after trypsin digest (data not shown).

NPro Adopts a Cadherin-like Fold

The NPro structured region (residues 31–124) has a Greek-key topology of a β sandwich-like fold, consisting of seven β strands (A–G, Figure 3B). All strands are antiparallel except for the parallel pairing between β-A and β-G. β-A, -G, -F, and -C form one sheet of the β sandwich, and β-B, -E, and -D form the second sheet. A third small sheet is formed by β-A′ and β-B′. A disulfide bridge on top of the β sandwich stabilizes the β-A′ and β-B′ pairing (Figure 3C). The two short β strands were named as extensions of β-A and β-B for better comparison with cadherin domain structures (see below). Prosequences from other cadherins (in particular E-, C-, and R-cadherin) most likely have a very similar fold because of significant sequence similarities to NPro within the structured region (Figure 1B).

Prosequences of classical cadherins are considerably longer than cadherin domain sequences, and our structural analysis reveals that this is largely due to an unstructured C-terminal region (Figure 4A). Despite insignificantly low sequence similarities between NPro and cadherin domains, the topology of the folded region of NPro is remarkably similar to that of cadherin domains; position and length of the seven main β strands are all well conserved, and several structurally important, core-defining residues (e.g., Val58, Phe79, Gly85) are not only conserved among prodomains but also between prodomains and cadherin domains (Figures 4A and 1B). Small but potentially crucial variations between NPro fold and cadherin fold can be found mainly in the loop or helical regions between the seven main β strands (Figure 4B). In particular, NPro has no α helix between β-B and β-C and significantly shorter C/D and F/G loops. It is remarkable that even in a secondary structure-guided alignment, the sequence similarity between NPro and CAD1s is very low (Figure 4A). Therefore, we conclude that NPro—and most likely other prodomains too—has a

Table 1. NMR Structural Statistics for the N-Cadherin Prodomain

Total experimental restraints	2129	
Total NOE distance restraints	1955	
Ambiguous	190	
Unambiguous	1765	
Manually assigned	1217	
ARIA assigned	548	
Intraresidue	841	
Interresidue	924	
Sequential $\|i - j\| = 1$	334	
Medium $2 \leq \|i - j\| \leq 4$	82	
Long range $\|I - j\| > 4$	508	
Hydrogen bond restraints	60	
Dihedral angle restraints	114	
Final energies (kcal/mol)[a]		
E_{TOT}	223.7 ± 6.7	
E_{NOE}	19.1 ± 3.3	
E_{DIH}	1.4 ± 0.4	
E_{LJ}	−552.1 ± 19.6	
Ramachandran Plot (%)[a]	Full Molecule[b]	Secondary Structure[c]
Most favorable regions	78.2 ± 2.0	96.9 ± 2.2
Additional allowed regions	15.6 ± 2.4	3.1 ± 2.2
Generously allowed regions	4.6 ± 1.5	0.0 ± 0.0
Disallowed regions	1.6 ± 0.8	0.0 ± 0.0
Rmsds of atomic coordinates (Å)[a]		
Backbone	0.67 ± 0.08	0.37 ± 0.06
Heavy atoms	1.29 ± 0.08	0.85 ± 0.07

[a] Based on the 20 lowest energy structures. None of these final structures exhibit NOE-derived distance violations greater than 0.5 Å or dihedral angle restraint violations greater than 5°.
[b] Residues 34–122.
[c] Residues 34–36, 40–44, 55–57, 60–62, 69–73, 79–82, 86–89, 99–108, 113–122.

cadherin-like fold that could not have been predicted from the primary sequence.

Structural Differences between NPro and Cadherin "Adhesive" Domains

We further analyzed the structure and sequence of NPro in comparison to cadherin domains in order to determine the structural basis for their functional differences. Adhesive interactions between cadherin domains require Ca^{2+} ions at millimolar concentrations (Koch et al., 1997; Pertz et al., 1999), and calcium stabilizes the quaternary structure of cadherins (Nagar et al., 1996; Pokutta et al., 1994). Furthermore, electron microscopy studies of recombinantly oligomerized cadherin ectodomains (Ahrens et al., 2002; Pertz et al., 1999) and NMR studies of ECAD12 (Haussinger et al., 2002) demonstrated that Ca^{2+} binding might also influence *cis*- and *trans*-interactions. None of the highly conserved calcium-coordinating residues in cadherin domains, such as Glu11, Asp67, and Glu69, are present in NPro (Figure 4A). In addition, instead of a cluster of Ca^{2+} binding residues at the C-terminal region, NPro has a much longer unstructured loop followed by a conserved protease cleavage site.

Cadherin-based cell adhesion is generated by homophilic *cis*- and *trans*-interactions between N-terminal, "adhesive" cadherin domains, and several, in part contradicting, models have been proposed for the molecular basis of such interactions (Boggon et al., 2002; Koch et al., 1999; Leckband, 2002; Troyanovsky, 1999). The significance of a conserved tryptophan residue (Trp2) for interactions has been demonstrated, although different modes of interactions have been proposed. The side chain of Trp2 was often found to be inserted into a conserved, hydrophobic pocket, and both structural and mutational data have revealed the importance of Trp2 and its acceptor pocket for adhesion (Chitaev and Troyanovsky, 1998; Kitagawa et al., 2000; Pertz et al., 1999; Shan et al., 2000; Tamura et al., 1998; Tanaka et al., 2000). Originally, Trp2 was thought to be involved in *cis*-interactions (Shapiro et al., 1995), but in a recent C-cadherin structure, which includes all five extracellular domains (ectodomain), the same Trp-mediated interface was interpreted as *trans*-interaction, because the curved ectodomains are clearly arranged in opposite directions (Boggon et al., 2002). Very recently, it has been proposed that Trp2 can be engaged in both *cis*- and *trans*-interactions, reconciling earlier observations (He et al., 2003).

Notably, NPro has a hydrophobic pocket similar to that found in cadherin domains (Figure 4C). However, CAD1s have two extra residues between β-F and the shorter β-G strand, resulting in a bulge within the F-G loop (Figure 4B) and a hydrophobic cavity that is large enough to accommodate the side chain of a conserved, functionally important Trp, whereas NPro has a Phe side chain (Phe36) docked into a smaller hydrophobic pocket (Figure 4C). In this respect NPro is more similar to NCAD2—and other CAD2s—having a smaller hydrophobic cavity with the side chain of a Phe inserted instead of a Trp (Figure 4C). Unlike Trp2 in CAD1s, Phe36 in NPro is not part of a flexible strand and is essentially immobilized in its position within the hydrophobic core

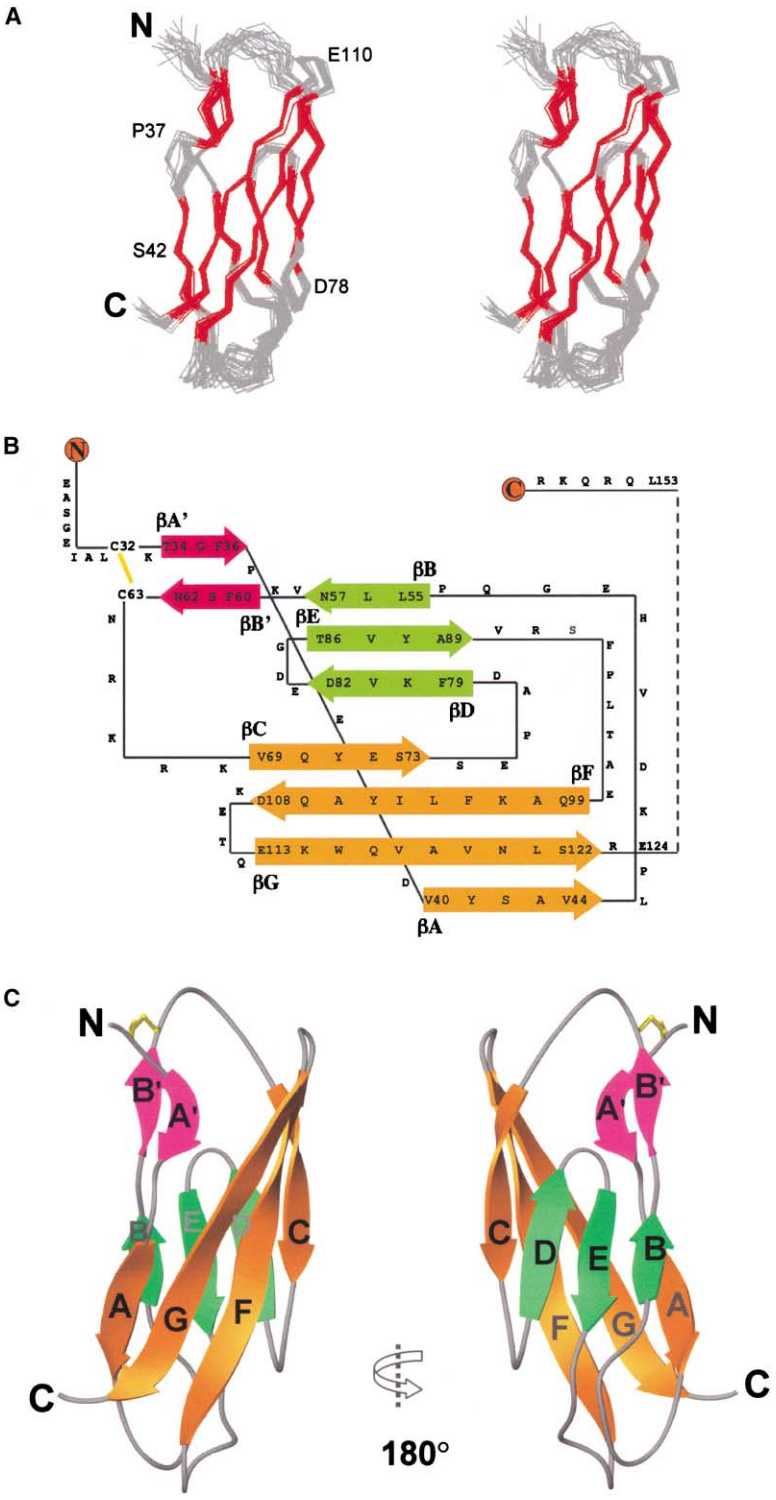

Figure 3. Overview of the NPro Structure

(A) Superposition (N-, C_α-, and C′-atoms) of the 20 lowest energy NMR-derived structures of NPro (residues 31–124) in stereo view. Structurally disordered terminal regions were omitted for clarity, and β strand regions are colored in orange.

(B) Greek-key type topology of NPro. β strands are shown as magenta, green, or orange arrows representing three different sheets. A disulfide bond is depicted in yellow. Residue numbers indicate the boundaries of the strands, and the single-letter code for amino acids was used. N and C termini are depicted by spheres. Amino acid residues at the unstructured C terminus (Pro 125–Arg159) were omitted (indicated as a dashed line); only residues of the proteolytic cleavage site are shown.

(C) Ribbon diagram of a representative low-energy NMR structure of NPro (residues 31–124). β strands of the three sheets are depicted as magenta, green, and orange arrows, and the disulfide bridge between Cys32 and Cys63 is depicted in yellow. Loop regions are shown in gray, and N- and C termini are indicated.

of a β sheet, which gains additional stability from a disulfide bridge (Figure 4C). Therefore, it is unlikely that this residue could be involved in *cis*- or *trans*-interactions. Correspondingly, we have no experimental evidence for NPro homophilic interactions. Even at the high protein concentrations (~0.8 mM) used for NMR experiments, we did not observe dimer formation or any form of aggregation, as judged by gel filtration chromatography or ^{15}N-HSQC spectral analysis (data not shown).

e also analyzed the electrostatic surface potential of NPro in order to determine if, despite differences in the sequence, NPro still has potential interaction sites similar to those found in cadherin domains. We utilized the structural model derived from the high-resolution C-cadherin structure (Boggon et al., 2002) to compare candidate *trans*- and *cis*-interaction sites with corresponding sites in NPro. The *trans*-interaction site in C-cadherin (Boggon et al., 2002) is very similar to that

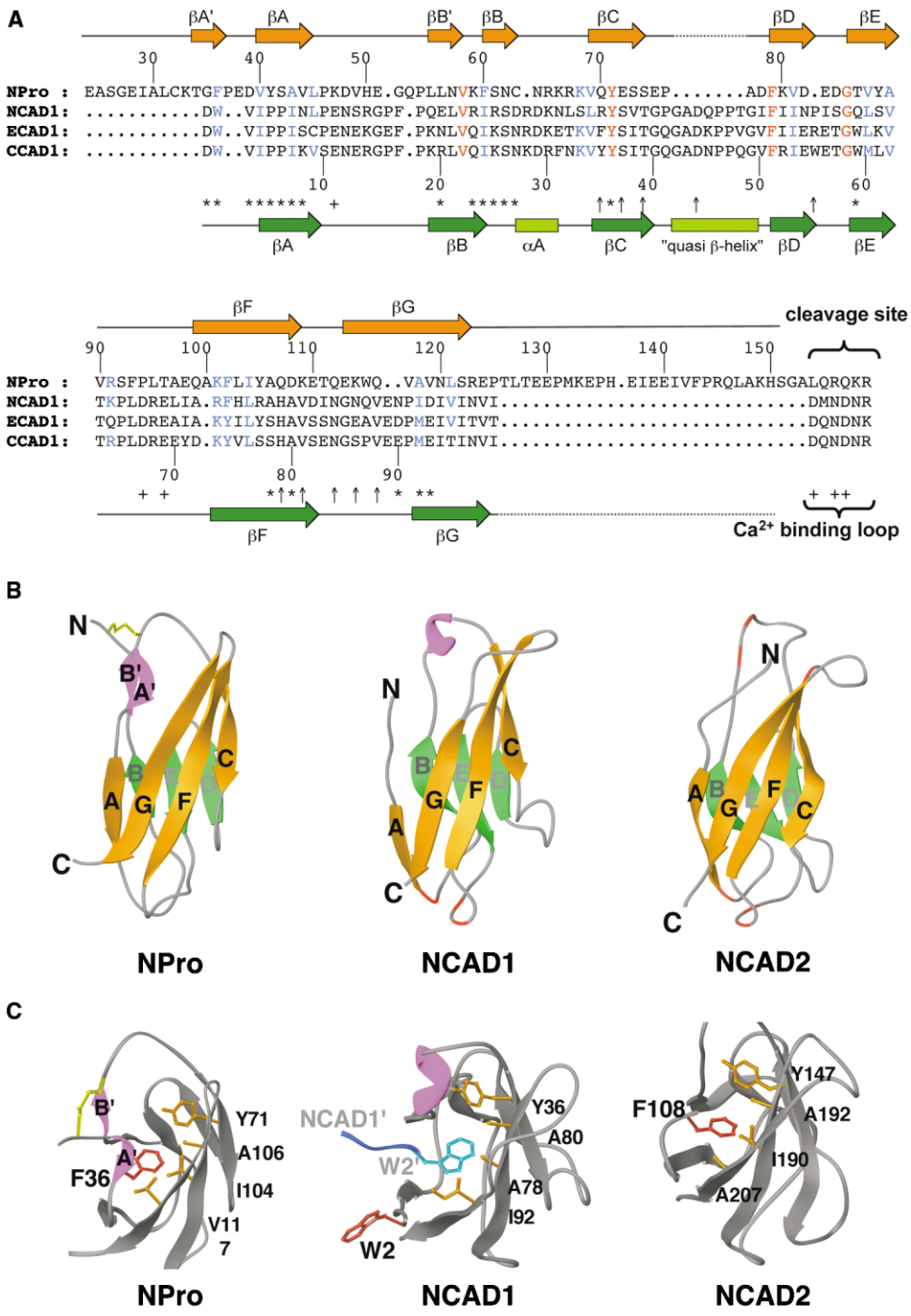

Figure 4. Comparison of NPro and Cadherin Domains

(A) Structure-based sequence homology alignment of NPro with N-terminal cadherin domains. The N-cadherin prodomain sequence (NPro) and three N-terminal cadherin domain (CAD1) sequences were aligned. Residue numbers for NPro (above aligned sequences) are according to the annotation for full-length cadherin sequences as found in databases. CAD1 residue numbers (below aligned sequences) reflect the mature cadherin proteins without prosequences and signal peptides. The positions of β strands from the NPro structure are indicated by orange arrows above the aligned sequences. Secondary structure elements from CAD1 structures are shown as light green rectangles (α helix and "quasi β helix") and green arrows (β strands) underneath the aligned sequences. Absolutely conserved, identical amino acids are colored red, and highly conserved, similar amino acids are colored blue. In addition, conserved Ca^{2+} binding residues in CAD1s are indicated by crosses (+); residues involved in *trans*-interactions in C-cadherin by stars (*); and residues involved in *cis*-interaction in C-cadherin by arrows (↑) (Boggon et al., 2002). All sequences are from mouse except C-cadherin, which is from *Xenopus laevis*.

(B) Ribbon representation of the NPro structure (left), the CAD1 structure of N-cadherin (PDB accession codes 1NCG and 1NCI, middle), and the CAD2 structure of N-cadherin (PDB accession code 1NCJ, right). The same color-coding is used as in Figure 3C. Additionally, the positions of Ca^{2+} binding residues are indicated in red.

(C) Comparison of the conserved hydrophobic pocket in NPro and in three cadherin domain structures. NPro: hydrophobic pocket of NPro, in which Phe36 inserts its side chain; NCAD1: hydrophobic pocket in CAD1 of N-cadherin with the buried Trp2 side chain (W2′) that projects from the equivalent partner molecule (NCAD1′) of a dimer (intermolecular interaction); NCAD2: hydrophobic pocket in CAD2 of N-cadherin with the buried Phe108 side chain. The single-letter code is used for amino acids, and the side chains are shown in stick representation. Residue numbers are according to (A).

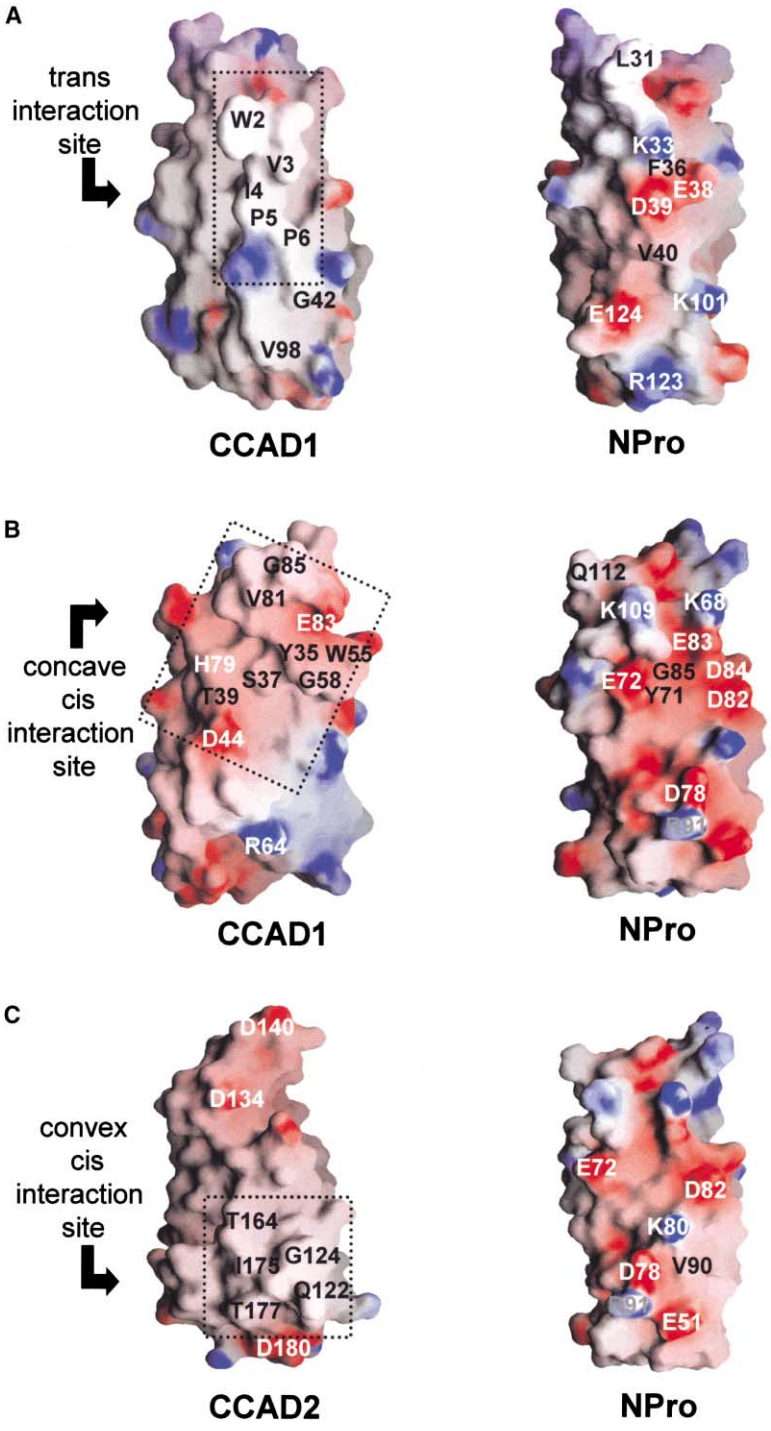

Figure 5. Comparison of Electrostatic Potential Surface Maps between C-Cadherin and NPro

Negatively and positively charged residues are shown in red and blue, respectively. Black and white numerals indicate the location of hydrophobic and charged residues, respectively. The single-letter code is used for amino acids.

(A) Left: *trans*-interaction site as seen in the CAD1 domain of C-cadherin (PDB accession code 1L3W). Some of the residues that are involved in *trans*-interaction and that are conserved among CAD1s are indicated (W2-P6). Note that the side chain of Trp2, seen here, is buried into the hydrophobic pocket of a partner molecule that is omitted here for clarity. Right: corresponding surface in NPro. Some of the negatively charged residues seen on this surface are indicated. Note that Phe36 here indicates the backbone of this phenylalanine and that its side chain is buried into the hydrophobic core of the same molecule.

(B) Left: concave *cis*-interaction site as seen in the CAD1 domain of C-cadherin. Some of the residues involved in *cis*-interactions are indicated (see also Figure 4A). Right: corresponding surface in NPro.

(C) Left: convex cis-interaction site as seen in the CAD2 domain of C-cadherin. Some of the residues involved in *cis*-interactions are indicated (Boggon et al., 2002). Right: corresponding surface in NPro.

observed in an earlier N-cadherin structure (Shapiro et al., 1995), only that it was then interpreted as the *cis*-interaction site. In contrast to the large hydrophobic *trans*-interaction surface of CCAD1, strikingly, NPro has a surface covered with highly negatively charged residues, including Glu38 and Asp39, whereas the side chains of hydrophobic residues, such as Phe36, are completely buried (Figure 5A).

In addition to Trp-mediated *trans*-interactions, *cis*-interactions were described between a concave surface in CAD1 and a complementary convex surface in CAD2 in C-cadherin (Boggon et al., 2002). The *cis*-interaction site in CAD1s includes a quasi-β helix between β-C and β-D that is absent in NPro (Figures 4A and 4B). Additionally, the surface potential of the concave *cis*-interaction site in C-cadherin, bordered by Gly85 and Asp44, is very different from a corresponding surface in NPro (Figure 5B). NPro has a stretch of largely negatively charged residues in this area (Asp82, Glu83, Asp84), and all the conserved *cis*-interaction residues in CAD1s

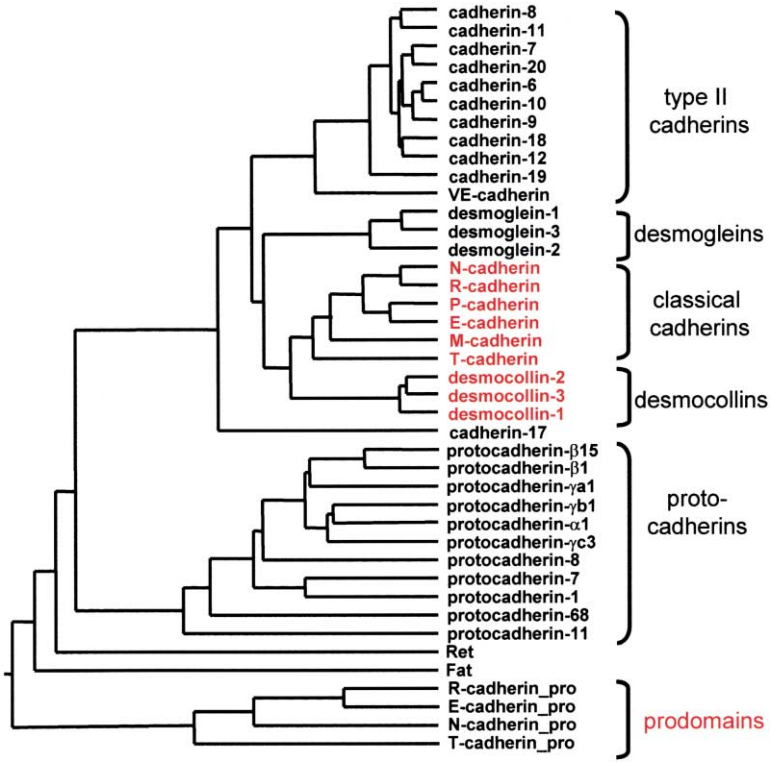

Figure 6. Phylogenetic Tree of Cadherin Domains and Prodomains

A phylogenetic tree was constructed with N-terminal cadherin domains and cadherin prodomains assuming an evolutionary clock. All sequences are from human. The names of the subfamilies are indicated, and prodomain-containing cadherins are colored red.

are not present in NPro or other prodomains (Figure 4A). In C-cadherin, hydrophobic residues from the lower part of the β-B-D-E sheet in CAD2 form a convex *cis*-interaction site. NPro again displays mainly charged residues on an equivalent surface, and no convex interaction site is present (Figure 5C).

Residues that are conserved between prodomains and CAD1s (Figure 4A) are all core defining residues and therefore important for the structure itself rather than for homophilic interactions. Mutations of some of the corresponding residues in R-cadherin (Tyr36, Phe51, Gly58) were shown to abolish adhesion (Kitagawa et al., 2000). However, in light of both the different cadherin structures and our NPro structure, the loss of adhesive capacity in these mutants should be interpreted as the result of a compromised structural integrity and not as a consequence of disrupting specific cadherin interactions.

Prodomains and Evolution of Cadherin Domains

To analyze the relationship between cadherins with prosequences of different length and between cadherin domains and prodomains, we performed a phylogenetic analysis of cadherin domain sequences and prodomain sequences from all the major cadherin subfamilies. A sequence alignment of 36 N-terminal cadherin domains and the structured region of four prodomains (all from human; data not shown) were used for phylogenetic analysis (Figure 6). In this simulated evolution, prodomains clearly diverge early from all other cadherin domains and can therefore be regarded as distant relatives of cadherin domains, with which they share a similar

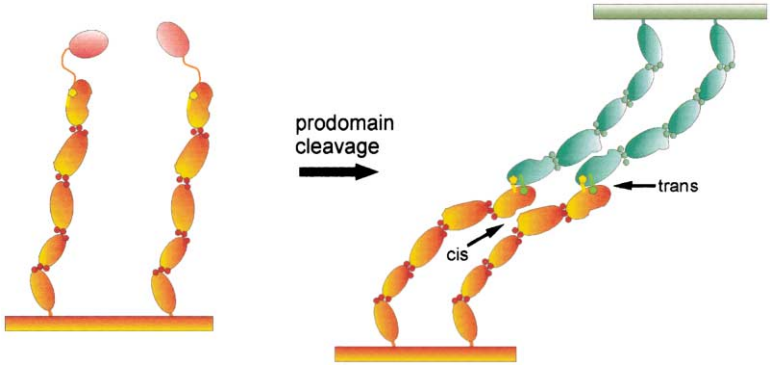

Figure 7. Model for Activation of Classical Cadherins through Prodomain Cleavage

Prodomain-bearing cadherins (cadherins with additional "nonadhesive" domains) do not interact with each other (left). After prodomain cleavage—e.g., during maturation in the Golgi—cadherins are activated to form homophilic *cis*-interactions between molecules from the same cell surface and homophilic *trans*-interactions between molecules from opposing cell surfaces (right). Cadherin domains are depicted as orange and green ovals, with the different colors symbolizing cadherins protruding from opposite cell surfaces. Prodomains are depicted as red ovals. Ca^{2+} ions are shown as red or green spheres, and Trp2 side chains are shown as yellow or green pentagons. *Cis*- and *trans* interfaces of activated cadherins are indicated and based on a recent C-cadherin crystal structure (Boggon et al., 2002).

folding topology. Concerning cadherins, our evolutionary analysis of cadherin domains and prodomains led to similar conclusions as reported previously for just the cadherin domains (Nollet et al., 2000). But notably, the prodomain-containing cadherins are clustered together, and desmocollins are more closely related to classical cadherins than to desmogleins (Figure 6).

Implications for Cell Adhesion Mechanisms

Our data clearly demonstrate that all the structural features known to be important for homophilic cadherin interactions are absent in NPro, providing a structural basis for the inability of NPro, and most likely other prodomains as well, to interact either with themselves or with cadherin domains. This is also supported by the fact that we were unable to detect any interactions between NPro and recombinant purified N-terminal cadherin domains at millimolar concentrations by NMR (data not shown). Recently, Ozawa had shown that the presence of the prosequence in E-cadherin indeed prevents dimer formation (Ozawa, 2002), supporting our structure-based prediction. Together, this shows that prodomains have no homophilic interaction capacities. Given the importance of intricate *cis*- and *trans*-interactions between N-terminal "adhesive" domains for cadherin-mediated adhesion, we can speculate that the presence of an extra "nonadhesive" domain close to "adhesive" domains, together with the rotational freedom provided by a long flexible C-terminal loop, might render these interactions sites less accessible and might also prevent a propagation of interactions. This explains why prodomain-bearing cadherins are unable to interact with each other and are therefore "nonadhesive" (Figure 7).

Very recently, electron tomography of skin sections has been used to unravel interactions between desmosomal cadherins (He et al., 2003). In this study, which for the first time provides a structural view of an entire cadherin-based junction in situ, C-cadherin crystal structure data were used to assign densities obtained from 3D reconstructions of electron tomography experiments to individual molecules. This work showed that desmosomal cadherins form clusters of four to six molecules that interact primarily via their N-terminal domains and that Trp2 can be engaged in both *cis*- and *trans*-interactions within these clusters. The presence of even only one or two prodomains would effectively prevent the formation of such clusters, just by steric hindrance. There is also an alternative model for cadherin interactions suggesting that multiple cadherin domains are involved in forming the adhesive interface (Chappuis-Flament et al., 2001; Sivasankar et al., 2001). Also in this model, the presence of "nonadhesive" prodomains would obstruct such a precise array of domain-domain interactions.

Incomplete cleavage of the E-cadherin prosequence on the cell surface did not restore cell adhesion, indicating that a few extra residues at the N-terminal part of mature E-cadherin can have an inhibitory effect (Ozawa and Kemler, 1990). This notion has also been brought forward in crystallographic studies in order to explain the different interactions seen in X-ray structures of different N-terminal cadherin fragments (Koch et al., 1999; Tamura et al., 1998). While a few extra residues might already be sufficient to abolish *cis*- and/or *trans*-interactions on the cell surface, the in vivo function of cadherin prodomains seems to be to prevent homophilic interactions during biosynthesis, especially within the Golgi (Ozawa, 2002; Ozawa and Kemler, 1990; Wahl et al., 2003). Prosequence cleavage is the last step in cadherin biosynthesis and activates cadherins just before they are transported to the cell surface (Wahl et al., 2003). Since our structural data help to explain why cadherins with uncleaved prodomains are unable to mediate adhesion, this work also provides the first structural glimpses into the largely unknown but potentially important process of cadherin activation by prodomain removal.

Experimental Procedures

Protein Expression and Purification

The DNA sequence, encoding the prodomain of N-cadherin (residues Glu24–Arg159), was amplified by PCR from mouse N-cadherin cDNA and subcloned into a pET-19b vector (Novagen). The resulting plasmid expresses the recombinant prodomain, containing an N-terminal histidine tag followed by an enterokinase cleavage site, in *E. coli* BL21 (DE3) cells (Novagen) grown in LB medium. Protein expression was induced at 25°C by adding 0.1 mM IPTG. Soluble proteins were purified by Ni^{2+}-NTA affinity chromatography (Qiagen), followed by extensive dialysis and gel filtration chromatography (Superose 12/20, Amersham) on FPLC (Pharmacia). The histidine tag was cleaved by enterokinase treatment (in 20 mM Tris, 140 mM NaCl [pH 7.4]) and removed by rechromatography on Ni^{2+}-NTA. Correct cleavage was confirmed by mass spectrometry and N-terminal sequencing (Protein Core Facility, Columbia University, New York). The recombinant proteins appeared more than 95% pure as judged by SDS-PAGE and mass spectrometry. For disulfide bond analysis, NPro was digested with trypsin at room temperature overnight in the absence and presence of DTT (10 mM), and the resulting proteolytic fragments were analyzed by MALDI (Protein Core Facility, Columbia University, New York).

NMR Sample Preparation

Uniformly ^{15}N- and $^{15}N/^{13}C$-labeled NPro samples were prepared by growing *E. coli* BL21 (DE3) cells in M9 minimal medium containing $^{15}NH_4Cl$ (1 g/l) with or without $^{13}C_6$-glucose (2 g/l). Uniformly $^{15}N/^{13}C$-labeled and fractionally deuterated protein samples were prepared by growing cells in 75% ^2H_2O. NPro samples were purified as described above, but the histidine tag was not removed for NMR samples. After gel filtration chromatography, protein samples were concentrated and dialyzed. The final NMR samples were typically 0.8–1 mM protein in 50 mM sodium phosphate buffer (pH 6.0), containing 0.05% sodium azide (w/v), in $H_2O/^2H_2O$ (9/1) or in ^2H_2O. The proteins appeared pure on SDS-PAGE, and the quality of the samples was assessed by ^{15}N-HSQC spectra.

NMR Spectroscopy

All NMR spectra were collected at 30°C on a Bruker DRX 600 or 500 MHz spectrometer. The backbone and side chain ^1H, ^{13}C, and ^{15}N resonances of the protein were sequentially assigned using deuterium-decoupled triple resonance spectra of HNCACB, HN(CO)CACB, and (H)C (CO)NH-TOCSY recorded on uniformly $^{15}N/^{13}C$-labeled and fractionally deuterated protein samples (Sattler et al., 1999; Yamazaki et al., 1994). To facilitate backbone assignment, ^{15}N-HSQC spectra of protein samples prepared from media with selectively labeled amino acids (^{15}N-Leu, -Val, -Glu, or -Phe) were also recorded. Side chain assignments were completed with 3D HCCH-TOCSY spectra collected from a uniformly $^{15}N/^{13}C$-labeled protein sample (Clore and Gronenborn, 1994). NOE-derived distance restraints were obtained from ^{15}N- or ^{13}C-edited 3D NOESY spectra (Clore and Gronenborn, 1994). All NMR spectra were processed with NMRPipe/

NMRDraw (Delaglio et al., 1995) and analyzed with NMRView (Johnson and Blevins, 1994).

Structure Calculation

Structures of NPro were calculated with a distance geometry-simulated annealing protocol using the X-PLOR program (Brunger, 1993). Initial structure calculations were performed with manually assigned NOE-derived distance restraints. Hydrogen bond distance restraints were then added for residues with characteristic NOE patterns. Finally, ARIA, an iterative automated assignment program that also integrates with X-PLOR, was used for structure refinement (Nilges and O'Donoghue, 1998). The final structure calculation used a total of 2129 experimental restraints, of which 1955 were NOE-derived distance restraints obtained from manual and ARIA-assisted assignments of ^{15}N- and ^{13}C-edited NOE data. The NOE-derived restraints were categorized based on the observed NOE peak intensities. A total of 60 hydrogen bond restraints and 114 dihedral angle restraints, obtained from chemical shift index analysis employing the TALOS program (Cornilescu et al., 1999), were also used in the calculations. PROCHECK (Laskowski et al., 1996) was used to validate an ensemble of 20 final NMR structures. Figures were generated using the programs InsightII (Molecular Simulations), Ribbons (Carson, 1991), and Grasp (Nicholls et al., 1993).

Alignments and Phylogenetic Tree Construction

The program ClustalX (version 1.8) was used for aligning protein sequences (Thompson et al., 1997). A distant matrix phylogenetic tree was constructed for 37 cadherin domains and 4 cadherin prodomains using programs from PHYLIP (Phylogeny Inference Package) version 3.5c (Felsenstein, 1989). In particular, the KITSCH algorithm was applied to build a phylogenetic tree assuming an evolutionary clock. The program TreeEdit version 1.0a (http://evolve.zoo.ox.ac.uk/software/TreeEdit/main.html) was used to display the phylogenetic tree. The human cadherin sequences used for alignment and subsequent phylogenetic tree construction were all retrieved from GenBank.

Polyclonal Antibody Generation

Polyclonal antibodies directed against the entire prodomain of N-cadherin were generated in rabbits (Covance Research Products, Denver, PA). Purified NPro (see above) was used at a concentration of 0.5 mg/ml in a 20 mM Tris buffer (pH 7.4), containing 150 mM NaCl. Antisera were tested by immunoblotting of cell extracts from L cells transfected with N-cadherin cDNA or from untransfected cells (see below). For immunostaining, affinity-purified NPro antibodies were used. An antigen affinity column was prepared by coupling purified recombinant NPro to Affi-Gel 15 (Bio-Rad). The coupling procedure and preparation of the antigen column were done according to the manufacturer's suggestions (Bio-Rad), and 10–12 mg of antigen was coupled to 0.5 ml Affi-Gel 15 resin. IgG antibody fractions were isolated from crude sera with the Affi-Gel Protein A MAPSII kit (Bio-Rad), dialyzed into 10 mM Tris (pH 7.5), and applied to the antigen affinity column. After washing, the affinity-purified IgG antibodies were eluted with 100 mM glycine (pH 2.5), immediately neutralized, dialyzed into a 10 mM Tris buffer (pH 8), and concentrated (Centricon concentrators, Amicon).

Cell Culture, Mutagenesis, and Transfection

L cells were cultured in DMEM supplemented with 10% FCS, 100 U/ml penicillin, and 100 mg/ml streptomycin. For mutagenesis, a 1.1 kb fragment containing the prodomain region was cut out and inserted into the PstI site of the pGEM-3zf (−) vector (Promega). Mutagenesis was carried out using the QuikChange site-directed mutagenesis kit (Stratagene). To obtain wild-type and mutant N-cadherin-expressing cells, cDNAs inserted into $pCXN_2$ vectors were transfected into mouse L cells by Superfect (Qiagen). The day after transfection, cells were split and seeded in complete DMEM containing 800 μg/ml of Geneticin G418 (GIBCO) to select stable transfected cell lines.

Immunocytochemistry and Cell Aggregation Assays

Immunocytochemistry was performed as described before (Shan et al., 2000). Our affinity-purified rabbit anti-NPro antibodies (see above) and rat anti-EC2 N-cadherin antibodies were used as primary antibodies. After incubation for 1 hr at 37°C with primary antibodies (rabbit NPro and rat NCAD2 antibodies), cells were washed and then incubated with fluorescent-conjugated secondary antibodies (Jackson ImmunoResearch Laboratories) at room temperature for 30 min. Coverslips were then mounted and examined by confocal laser microscopy.

For cell aggregation assays, monolayer cultures were treated with 2 mM EGTA in HCMF (HEPES-buffered Ca^{2+}- and Mg^{2+}-free Hanks' solution) for 30 min at 37°C. The trypsinized (ProN-trypsin cell line was with 0.01% trypsin and 1 mM $CaCl_2$) and nontrypsinized cell lines were washed gently in HCMF containing calcium and 1% BSA at 4°C. After the cells were thoroughly dissociated, 5×10^5 cells per well were transferred to 24-well dishes for a final volume of 0.5 ml HCMF, containing 1% BSA with 1 mM Ca^{2+}. The plates were rotated at 80 rpm at 37°C. The aggregates were observed in the time course. Factor Xa (0.2 U/ml) was added before and after cell dissociation.

Acknowledgments

We wish to thank L. Shapiro for his support during the initial phase of this work. We are grateful to J. Engel, T. Ahrens, and K. Manzur for critical reading of the manuscript. We are also grateful to M.A. Gawinowicz for performing mass spectrometry analysis. This work was supported by grants from the National Institute of Health to M.-M.Z. (CA80938) and to D.R.C. (NS41687), and a New York State Spinal Cord Injury Trust Award to D.R.C. (CO17683). A.W.K. is a recipient of a Swiss National Science Foundation Postdoctoral Fellowship, and A.F. is a recipient of a Welcome Trust Postdoctoral Fellowship. D.R.C. is a recipient of a Canada Research Chair and is also supported by the Canada Institutes of Health Research and the Canada Foundation for Innovation.

Received: January 16, 2004
Revised: February 18, 2004
Accepted: February 18, 2004
Published: May 11, 2004

References

Aberle, H., Schwartz, H., and Kemler, R. (1996). Cadherin-catenin complex: protein interactions and their implications for cadherin function. J. Cell. Biochem. *61*, 514–523.

Ahrens, T., Pertz, O., Haussinger, D., Fauser, C., Schulthess, T., and Engel, J. (2002). Analysis of heterophilic and homophilic interactions of cadherins using the c-Jun/c-Fos dimerization domains. J. Biol. Chem. *277*, 19455–19460.

Boggon, T.J., Murray, J., Chappuis-Flament, S., Wong, E., Gumbiner, B.M., and Shapiro, L. (2002). C-cadherin ectodomain structure and implications for cell adhesion mechanisms. Science *296*, 1308–1313.

Brieher, W.M., Yap, A.S., and Gumbiner, B.M. (1996). Lateral dimerization is required for the homophilic binding activity of C-cadherin. J. Cell Biol. *135*, 487–496.

Brunger, A.T. (1993). X-PLOR Version 3.1: A System for X-Ray Crystallography and NMR, Version 3.1 Edition (New Haven, CT: Yale University Press).

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958–961.

Cavallaro, U., Schaffhauser, B., and Christofori, G. (2002). Cadherins and the tumour progression: is it all in a switch? Cancer Lett. *176*, 123–128.

Chappuis-Flament, S., Wong, E., Hicks, L.D., Kay, C.M., and Gumbiner, B.M. (2001). Multiple cadherin extracellular repeats mediate homophilic binding and adhesion. J. Cell Biol. *154*, 231–243.

Chitaev, N.A., and Troyanovsky, S.M. (1998). Adhesive but not lateral E-cadherin complexes require calcium and catenins for their formation. J. Cell Biol. *142*, 837–846.

Clore, G.M., and Gronenborn, A.M. (1994). Multidimensional heteronuclear nuclear magnetic resonance of proteins. Methods Enzymol. *239*, 349–363.

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone

angle restraints from searching a database for chemical shift and sequence homology. J. Biomol. NMR *13*, 289–302.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.

Felsenstein, J. (1989). PHYLIP—Phylogeny Inference Package (Version 3.2). Cladistics *5*, 164–166.

Gumbiner, B.M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell *84*, 345–357.

Gumbiner, B.M. (2000). Regulation of cadherin adhesive activity. J. Cell Biol. *148*, 399–404.

Haussinger, D., Ahrens, T., Sass, H.J., Pertz, O., Engel, J., and Grzesiek, S. (2002). Calcium-dependent homoassociation of E-cadherin by NMR spectroscopy: changes in mobility, conformation and mapping of contact regions. J. Mol. Biol. *324*, 823–839.

He, W., Cowin, P., and Stokes, D.L. (2003). Untangling desmosomal knots with electron tomography. Science *302*, 109–113.

Huber, A.H., and Weis, W.I. (2001). The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. Cell *105*, 391–402.

Johnson, B.A., and Blevins, R.A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR *4*, 603–614.

Kemler, R. (1992). Classical cadherins. Semin. Cell Biol. *3*, 149–155.

Kitagawa, M., Natori, M., Murase, S., Hirano, S., Taketani, S., and Suzuki, S.T. (2000). Mutation analysis of cadherin-4 reveals amino acid residues of EC1 important for the structure and function. Biochem. Biophys. Res. Commun. *271*, 358–363.

Koch, A.W., Pokutta, S., Lustig, A., and Engel, J. (1997). Calcium binding and homoassociation of E-cadherin domains. Biochemistry *36*, 7697–7705.

Koch, A.W., Bozic, D., Pertz, O., and Engel, J. (1999). Homophilic adhesion by cadherins. Curr. Opin. Struct. Biol. *9*, 275–281.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR *8*, 477–486.

Leckband, D. (2002). The structure of the C-cadherin ectodomain resolved. Structure (Camb.) *10*, 739–740.

Lin, M.S., Mascaro, J.M., Jr., Liu, Z., Espana, A., and Diaz, L.A. (1997). The desmosome and hemidesmosome in cutaneous autoimmunity. Clin. Exp. Immunol. *107* (*Suppl 1*), 9–15.

Mengaud, J., Ohayon, H., Gounon, P., Mege, R.M., and Cossart, P. (1996). E-cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. Cell *84*, 923–932.

Nagar, B., Overduin, M., Ikura, M., and Rini, J.M. (1996). Structural basis of calcium-induced E-cadherin rigidification and dimerization. Nature *380*, 360–364.

Nicholls, A., Bharadwj, R., and Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. Biophys. J. *64*, 166–170.

Nilges, M., and O'Donoghue, S. (1998). Ambiguous NOEs and automated NOE assignment. Prog. NMR Spectrosc. *32*, 107–139.

Nollet, F., Kools, P., and van Roy, F. (2000). Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. J. Mol. Biol. *299*, 551–572.

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

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

Ozawa, M. (2002). Lateral dimerization of the E-cadherin extracellular domain is necessary but not sufficient for adhesive activity. J. Biol. Chem. *277*, 19600–19608.

Ozawa, M., and Kemler, R. (1990). Correct proteolytic cleavage is required for the cell adhesive function of uvomorulin. J. Cell Biol. *111*, 1645–1650.

Pertz, O., Bozic, D., Koch, A.W., Fauser, C., Brancaccio, A., and Engel, J. (1999). A new crystal structure, Ca2+ dependence and mutational analysis reveal molecular details of E-cadherin homoassociation. EMBO J. *18*, 1738–1747.

Pokutta, S., and Weis, W.I. (2002). The cytoplasmic face of cell contact sites. Curr. Opin. Struct. Biol. *12*, 255–262.

Pokutta, S., Herrenknecht, K., Kemler, R., and Engel, J. (1994). Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. Eur. J. Biochem. *223*, 1019–1026.

Posthaus, H., Dubois, C.M., Laprise, M.H., Grondin, F., Suter, M.M., and Muller, E. (1998). Proprotein cleavage of E-cadherin by furin in baculovirus over-expression system: potential role of other convertases in mammalian cells. FEBS Lett. *438*, 306–310.

Redies, C. (2000). Cadherins in the central nervous system. Prog. Neurobiol. *61*, 611–648.

Sansonetti, P.J., Mounier, J., Prevost, M.C., and Mege, R.M. (1994). Cadherin expression is required for the spread of *Shigella flexneri* between epithelial cells. Cell *76*, 829–839.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. NMR Spectrosc. *34*, 93–158.

Schubert, W.D., Urbanke, C., Ziehm, T., Beier, V., Machner, M.P., Domann, E., Wehland, J., Chakraborty, T., and Heinz, D.W. (2002). Structure of Internalin, a Major Invasion Protein of Listeria monocytogenes, in Complex with Its Human Receptor E-Cadherin. Cell *111*, 825–836.

Seidah, N.G., and Chretien, M. (1999). Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. Brain Res. *848*, 45–62.

Shan, W.S., Tanaka, H., Phillips, G.R., Arndt, K., Yoshida, M., Colman, D.R., and Shapiro, L. (2000). Functional cis-heterodimers of N- and R-cadherins. J. Cell Biol. *148*, 579–590.

Shapiro, L., Fannon, A.M., Kwong, P.D., Thompson, A., Lehmann, M.S., Grubel, G., Legrand, J.F., Als-Nielsen, J., Colman, D.R., and Hendrickson, W.A. (1995). Structural basis of cell-cell adhesion by cadherins. Nature *374*, 327–337.

Shimoyama, Y., Tsujimoto, G., Kitajima, M., and Natori, M. (2000). Identification of three human type-II classic cadherins and frequent heterophilic interactions between different subclasses of type-II classic cadherins. Biochem. J. *349*, 159–167.

Shore, E.M., and Nelson, W.J. (1991). Biosynthesis of the cell adhesion molecule uvomorulin (E-cadherin) in Madin-Darby canine kidney epithelial cells. J. Biol. Chem. *266*, 19672–19680.

Sivasankar, S., Gumbiner, B., and Leckband, D. (2001). Direct measurements of multiple adhesive alignments and unbinding trajectories between cadherin extracellular domains. Biophys. J. *80*, 1758–1768.

Takeda, H., Shimoyama, Y., Nagafuchi, A., and Hirohashi, S. (1999). E-cadherin functions as a cis-dimer at the cell-cell adhesive interface in vivo. Nat. Struct. Biol. *6*, 310–312.

Takeichi, M. (1988). The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. Development *102*, 639–655.

Takeichi, M. (1995). Morphogenetic roles of classic cadherins. Curr. Opin. Cell Biol. *7*, 619–627.

Tamura, K., Shan, W.S., Hendrickson, W.A., Colman, D.R., and Shapiro, L. (1998). Structure-function analysis of cell adhesion by neural (N-) cadherin. Neuron *20*, 1153–1163.

Tanaka, H., Shan, W., Phillips, G.R., Arndt, K., Bozdagi, O., Shapiro, L., Huntley, G.W., Benson, D.L., and Colman, D.R. (2000). Molecular modification of N-cadherin in response to synaptic activity. Neuron *25*, 93–107.

Tepass, U., Truong, K., Godt, D., Ikura, M., and Peifer, M. (2000). Cadherins in embryonic and neural morphogenesis. Nat. Rev. Mol. Cell Biol. *1*, 91–100.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and

Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876–4882.

Troyanovsky, S.M. (1999). Mechanism of cell-cell adhesion complex assembly. Curr. Opin. Cell Biol. *11*, 561–566.

Volk, T., Cohen, O., and Geiger, B. (1987). Formation of heterotypic adherens-type junctions between L–CAM-containing liver cells and A-CAM-containing lens cells. Cell *50*, 987–994.

Wahl, J.K., 3rd, Kim, Y.J., Cullen, J.M., Johnson, K.R., and Wheelock, M.J. (2003). N-cadherin-catenin complexes form prior to cleavage of the proregion and transport to the plasma membrane. J. Biol. Chem. *278*, 17269–17276.

Yamazaki, T., Lee, W., Arrowsmith, C.H., Mahandiram, D.R., and Kay, L.E. (1994). A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity. J. Am. Chem. Soc. *116*, 11655–11666.

Yap, A.S., Brieher, W.M., and Gumbiner, B.M. (1997a). Molecular and functional analysis of cadherin-based adherens junctions. Annu. Rev. Cell Dev. Biol. *13*, 119–146.

Yap, A.S., Brieher, W.M., Pruschy, M., and Gumbiner, B.M. (1997b). Lateral clustering of the adhesive ectodomain: a fundamental determinant of cadherin function. Curr. Biol. *7*, 308–315.

Accession Numbers

The atomic coordinates of the NPro structure have been deposited in the Protein Data Bank with the accession code 1OP4.