Molecular Modeling of the Extracellular Domain of the RET Receptor Tyrosine Kinase Reveals Multiple Cadherin-like Domains and a Calcium-binding Site*

Jonas Anders‡, Svend Kjær§, and Carlos F. Ibáñez¶

From the Division of Molecular Neurobiology, Department of Neuroscience, Karolinska Institute, 171 77 Stockholm, Sweden

Using bioinformatic tools, mutagenesis, and binding studies, we have investigated the structural organization of the extracellular region of the RET receptor tyrosine kinase, a functional receptor for glial cell line-derived neurotrophic factor (GDNF). Multiple sequence alignments of seven vertebrate sequences and one invertebrate RET sequence delineated four distinct N-terminal domains, each of about 110 residues, containing many of the consensus motifs of the cadherin fold. Based on these alignments and the crystal structures of epithelial and neural cadherins, we have generated molecular models of each of the four cadherin-like domains in the extracellular region of human RET. The modeled structures represent realistic models from both energetic and geometrical points of view and are consistent with previous observations gathered from biochemical analyses of the effects of Hirschsprung's disease mutations affecting the folding and stability of the RET molecule, as well as our own site-directed mutagenesis studies of RET cadherin-like domain 1. We have also investigated the role of Ca^{2+} in ligand binding by RET and found that Ca^{2+} ions are required for RET binding to GDNF but not for GDNF binding to the GFRα1 co-receptor. In agreement with these results, RET, but not GFRα1, was found to bind Ca^{2+} directly. Our results indicate that the overall architecture of the extracellular region of RET is more closely related to cadherins than previously thought. The models of the cadherin-like domains of human RET represent valuable tools with which to guide future site-directed mutagenesis studies aimed at identifying residues involved in ligand binding and receptor activation.

The RET receptor tyrosine kinase is an unusual receptor from many points of view. It cannot by itself bind its ligand, GDNF,[1] unless in a complex with another protein, the glycosyl phosphatidylinositol-anchored receptor GFRα1 (1, 2). In contrast with other receptor tyrosine kinases, there appears to be only one RET homologue in all species investigated so far. All members of the GDNF ligand family utilize RET as a signal transducing receptor subunit, with specificity being determined by cooperation between RET and different members of the GFRα family of glycosyl phosphatidylinositol-anchored receptors (1, 2). Both gain- and loss-of-function mutations in the *RET* gene have been identified in human diseases. Mutations in *RET* of patients with multiple endocrine neoplasias type 2A and 2B and familial medullary thyroid carcinoma induce constitutive activation of the RET tyrosine kinase and lead to congenital and sporadic cancers in neuroendocrine organs (3, 4). On the other hand, loss-of-function mutations in RET cause a dominant genetic disorder of neural crest development known as Hirschsprung's disease (HSCR), which results in the death of neurons in distal segments of the enteric nervous systems and colon aganglionosis (5).

The extracellular region of the RET molecule is peculiar compared with that found in other receptor tyrosine kinases in that it lacks leucine repeats, immunoglobulin, and fibronectin-like domains that are common in many other such receptors. The only recognizable motif within its over 600-residue-long extracellular region so far appears to be a stretch of 110 residues with similarity to members of the cadherin family of Ca^{2+}-dependent cell adhesion molecules (6, 7). Cadherins comprise a large and divergent superfamily with at least six subfamilies and several more divergent members, characterized on the basis of their domain architecture, genomic structure, and phylogenetic relationships (8). The extracellular region of cadherins is formed by a variable number of repeated modules (cadherin domains) of about 110 residues, all sharing a common consensus sequence and often with a Ca^{2+}-binding site present in between each of the domains. Several crystal structures of cadherin domains from epithelial and neural cadherins have been solved (9–12). Ca^{2+} binding appears to help to linearize and rigidify the structure and, at least according to one report, to promote dimerization of cadherin domains (12). Results from several laboratories have indicated that lateral dimerization or clustering of cadherins may increase their adhesivity (see, for example, Ref. 9). As originally proposed, the segment of similarity between RET and cadherins would extend from the middle of one cadherin domain to the next over a Ca^{2+}-binding motif (6). However, because the remaining portion of the RET extracellular region did not appear to display any obvious similarity to cadherins or to any other protein, the actual relationship between RET and cadherins has been unclear.

[1] The abbreviations used are: GDNF, glial cell line-derived neurotrophic factor; HSCR, Hirschsprung's disease; TBS, Tris-buffered saline; CLD, cadherin-like domain; E-cadherin, epithelial cadherin; N-cadherin, neural cadherin; Endo H, endoglycosidase H.

In the present study, we have used a battery of bioinformatic tools together with mutagenesis and binding studies to explore the structural organization of the extracellular region of RET. These studies allow us to propose a model for the overall architecture of the extracellular region of the RET molecule, as well as homology-based model structures of four of the five identified domains in this region.

MATERIALS AND METHODS

Sequence Analysis—Protein sequence data were retrieved from Gen-Bank™. Multiple sequence alignments were obtained by CLUSTALW running at the Decypher hardware accelerator (www.timelogic.com). For graphical presentation of multiple sequence alignments, the program TREEVIEW (13) was used. Sequence similarity searches were performed using PSI-BLAST running on the servers at either NCBI (www.ncbi.nlm.nih.gov/BLAST/) or at the Decypher hardware accelerator (www.timelogic.com/algo-blast2/PSIBLAST_aahtml-ssi). In both cases, the default settings of the program suggested by the input mask of the server were used with the low complexity filter being activated. Sequence-structure annotation was performed using the repertoire of vertebrate RET cadherin-like domains to search the Hidden Markov Model Library of protein domains implemented in the Structural Classification of Proteins data base release 1.53 available at stash.mrc-lmb.cam.ac.uk/SUPERFAMILY/hmm.html. Additionally, the fold of cadherin-like domains in the extracellular domain of RET was predicted using the three-dimensional position-specific scoring matrix method of enhanced genome annotation using structural profiles (14) available at www.bmm.icnet.uk/~3dpssm/. The program was used with the Global/Local and the low complexity filter option activated and with the sequences of vertebrate RET cadherin-like domains as baits.

Model Building—The Swiss-PdbViewer suite of programs (15) was used for comparative modeling of cadherin-like domains of human RET. Sequences of cadherin domains with three-dimensional structures available in the Brookhaven Protein Data Bank (16) were aligned with human RET cadherin-like domain sequences, superimposing the cadherin consensus sequence as identified in Fig. 2*A*. The hydrophobic residues represented in the cadherin consensus sequence were found to define the core of the classic cadherin domains with their side chains buried in the nonpolar interior of proteins (9, 11, 12). RET sequences were aligned to classic cadherin domains to optimize a match for consensus residues with an identical or conservative replacement. Gaps and insertions in the alignments were positioned within the variable loop regions in the cadherin template structures. As a constraint in model building, the superimposed cadherin consensus residues were fixed in position. The models were submitted to ProModII for model building using the Swiss-PdbViewer interface without adding any further constraint to the structure. The returned models were examined for unfavorable van der Waals' overlaps, which were removed by minimal rotation around side chain torsion angles. A short energy minimization was performed using the GROMOS 96 implementation in the Swiss-PdbViewer program suite (15).

The quality of the first models was evaluated manually and with a variety of computational techniques implemented in the World Wide Web-based Biotech structure and model verification server (biotech. ebi.ac.uk:8400/). Most quality indicators were in the normal range reported for computed modeled structures. As expected, some surface-located regions, especially those near gaps or insertions, were poorly modeled. The loops in the problematic regions were modeled individually, and the quality indicators for the model structures were then recalculated. When no suitable loop was available from a template cadherin structure, the loop data base implemented in the Swiss Pdb-Viewer program suite was searched to rank candidate loops from other proteins. The loop segment was then modeled on the best template that had no short contacts with the rest of the model.

Energy Minimization—The final models were energy minimized using the GROMOS 96 implementation in the Swiss-PdbViewer program suite (15). GROMOS computations were done *in vacuo* with the GROMOS 43B1 parameter set without reaction field. During the initial cycles of energy minimization, the α-carbon backbone was kept rigid, and the side chains alone were moved. Subsequently all atoms in the structures were allowed to move during energy minimization. This approach kept disturbances of the backbone structure to a minimum. Energy minimization was performed until all short contacts and inconsistencies in geometry were rectified. During the initial cycles of energy minimization, the electrostatic term was not included because the main object was to relieve steric clashes and to rectify bad geometry. The electrostatic term was then included in advanced stages of energy minimization.

Mutagenesis and Expression of RET Extracellular Domain Constructs—The extracellular domain of human RET, tagged with a hemagglutinin epitope, was subcloned in the pSectag2A vector (Invitrogen) for secretion of expressed proteins in mammalian cells. Point mutations were introduced using the QuickChange method (Stratagene) according to the instructions from the manufacturer. Cell lines stably expressing RET extracellular domain constructs were made by transfection of either human embryonic kidney (HEK293) or Chinese hamster ovary cells using the calcium phosphate transfection method. Hygromycin was applied as the selection marker. Individual clones were analyzed for protein expression by Western blotting. For protein expression, the cells were propagated in Dulbecco's modified Eagle's medium (Life Technologies, Inc.) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and 60 μg/ml gentamycin until 80% confluency. The cell monolayers were washed with phosphate-buffered saline and switched to serum-free medium for 96–120 h. Subsequently, the medium was harvested, the cell debris was removed by centrifugation, and the medium was concentrated by Amicon-10 ultrafiltration. For protein production in Chinese hamster ovary cells, sodium butyrate (2 mM) was included in the serum-free medium to enhance protein expression. This treatment resulted in substantial cell death after 3 days. Treatment with Endo H deglycosylase or Peptide:N-glycosidase F was performed according to the protocol supplied by the manufacturer (New England Biolabs). Deglycosylated proteins were subsequently separated by SDS-polyacrylamide gel electrophoresis, electroblotted to polyvinylidene difluoride membranes, and detected with an anti-hemagglutinin monoclonal antibody and anti-mouse secondary antibodies using chemiluminescence or chemifluorescence reagents from Pierce or Amersham Pharmacia Biotech, respectively.

Calcium Binding Assay—Microtiter plate wells (Nunc) were coated with 500 ng of an antibody recognizing the Fc portion of human immunoglobulin (Pierce). The wells were subsequently blocked with a Tris-buffered saline (TBS) solution containing 3% (w/v) bovine serum albumin for 1 h at room temperature. One hundred nanograms of mouse RET-Fc fusion protein (R&D Systems) in blocking solution was subsequently added to the wells. The same amount of GFRα1-Fc (R & D Systems) was added to control wells. Capture of the Fc fusion protein was allowed to proceed for 1 h at room temperature. Negative control wells, coated with only capturing antibody, were also included in the experiments. After rinsing the wells with TBS, 25 μM ^{45}Ca^{2+} (Amersham Pharmacia Biotech) in TBS/bovine serum albumin buffer was added to the wells and left for 1 h at room temperature. Subsequently, the wells were washed three times with TBS, bound ^{45}Ca^{2+} was released with 100 mM HCl, and the amount of radioactivity was counted. All data points were assayed in triplicate.

GDNF Radioiodination and Chemical Cross-linking—Binding studies were carried out with recombinant rat GDNF produced in Sf21 insect cells and purified as described previously (17). GDNF was iodinated by the lactoperoxidase method to a specific activity of $0.5–2 \times 10^8$ cpm/μg. ^{125}I-GDNF at 10 ng/ml was allowed to bind to cell monolayers at 4 °C during 3 h in binding buffer (50 mM HEPES, 137 mM NaCl) in the presence or the absence of Ca^{2+} ions (5 mM Cl$_2$Ca) or the Ca^{2+} chelator EGTA (at 5 mM) as indicated. After binding, cross-linking was initiated by addition of ethyl-dimethylaminopropyl carbodiimide supplemented with Sulfo-NHS (Pierce) or bis(succinimidyl) suberate as cross-linking agents for 30 min at 4 °C. Unlabeled GDNF was used at 50× molar excess to verify specific binding. Receptor complexes were fractionated by gradient SDS-polyacrylamide gel electrophoresis and visualized by autoradiography in a Storm 840 PhosphorImager (Molecular Dynamics).

RESULTS AND DISCUSSION

Sequence Alignment and Phylogeny of c-Ret Extracellular Domains—A GenBank™ search for RET sequences retrieved seven RET molecules from vertebrate species and one invertebrate RET from *Drosophila melanogaster*. The cytoplasmic kinase domain of RET is the best conserved region of the molecule, showing 90% sequence identity among the seven vertebrate species and 65% after including *Drosophila* RET. In the extracellular domain, a cysteine-rich region of ~120 residues containing a highly conserved pattern of 14 cysteine residues is found adjacent to the transmembrane segment. Although the cysteine pattern is highly conserved among all RET molecules in this segment, the residues between the cysteines
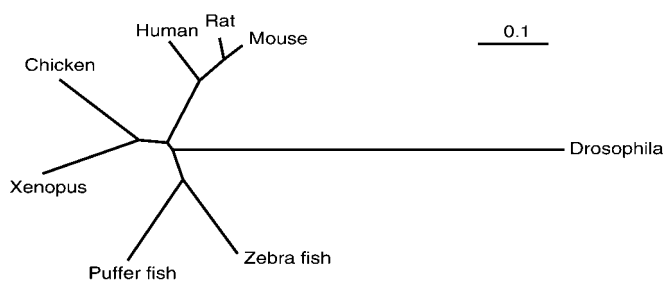
FIG. 1. **Phylogenetic tree of RET extracellular domains from different metazoan species.** Phylogenetic distances were based on the dissimilarity scores in the multiple sequence alignment. The graphical presentation of the unrooted tree was done with TREEVIEW (13).

are less well conserved, and the whole region shows 40% sequence identity among vertebrate RET molecules and 30% after including *Drosophila* RET. The conservation of RET extracellular sequences in the ~500 residues upstream of the cysteine-rich region is much lower, with only 20% identity among vertebrate RET sequences and 5% after including *Drosophila* RET. Included within this region, is a segment of about 110 residues previously identified as having sequence similarity to cadherins, containing a putative cadherin-like calcium-binding motif (6, 7). Despite the overall low sequence identity, the conservation of a cadherin-like calcium-binding motif and a juxtamembrane cysteine-rich region in all RET extracellular domains indicates that they are phylogenetically related (18). A phylogenetic tree of RET extracellular domains, derived from dissimilarity scores of pairwise sequence alignments, is shown in Fig. 1. The extracellular region of *Drosophila* RET is clearly most divergent, whereas vertebrate RET sequences cluster into three main groups comprising the fish, chicken/*Xenopus*, and mammalian RET sequences, respectively (Fig. 1). Data base searches of the *Drosophila* genome did not return any sequences with significant similarity to mammalian GFRα receptors or GDNF family ligands, molecules known to bind and activate RET in vertebrates (1).[2] The divergence of the *Drosophila* RET sequence upstream of the cysteine-rich domain might reflect a different function of the RET receptor in this organism.

Identification of Four Cadherin-like Domains in the Extracellular Region of RET—One of the most distinctive features of the extracellular region of RET is the segment with similarity to the calcium-binding motif of cadherins. This was initially identified as a 110-residue-long sequence from positions 203–314[3] in the extracellular region of vertebrate RET molecules displaying up to 40% sequence similarity with classical members of the cadherin family, including epithelial and neural cadherins. However, when mapped onto the crystal structures of epithelial and neural cadherin domains (9–12), the region of similarity to cadherins in RET does not correspond to a discrete cadherin but extends over the border of two consecutive domains including the intervening calcium-binding motif. This suggested that this segment of the RET molecule cannot represent a discrete structural domain of the protein and prompted us to extend the alignment of RET to cadherins both upstream and downstream of this region. Because of the repetitive organization of the ectodomain of cadherins, we hypothesized that the extracellular region of RET may also be organized by the repetitive arrangement of structurally related units.

We derived a consensus sequence for the cadherin domain from a multiple sequence alignment of the extracellular regions of the classical epithelial and neural cadherins, for which three-dimensional crystal structures are available, and protocadherin 7, a more distant member of the cadherin superfamily (19) (Fig. 2*A*). The consensus sequence derived from the comparison of the different cadherin domains in these three molecules is shown at the *bottom* of Fig. 2*A*. At the *top* of Fig. 2*A*, the seven prototypical strands of the cadherin fold are indicated (*arrows* mark *strands A–G*). Conserved features include a glutamic acid immediately C-terminal to strand A, a D*X*D motif C-terminal to strand B, a LDRE motif in the loop between strands E and F, and a negatively charged V*X*V*X*V*X*-D*X*ND*XX*P*X*F motif in strand G. The majority of cadherin domains also have V*X*Y*X*V (strand C), F*X*IE (strand D), and TG*X*L (strand E) motifs and a Y*X*L*XX*A*X*D motif C-terminal to strand F. Within the most N-terminal cadherin domain of classical cadherins, a single glutamic acid residue is conserved in the loop between strands F and G. Not only is the actual sequence of these motifs conserved among cadherins and cadherin-related proteins, but also the distances separating them are almost identical among different family members. This consensus sequence is conserved to a varying degree among more distantly related members of the cadherin superfamily (for a detailed analysis of the human repertoire of the cadherin superfamily see Ref. 8). As shown by the crystal structures of the first two domains of epithelial and neural cadherins (9–12), the cadherin domain adopts a β-sandwich fold, and calcium-binding sites are formed in between adjacent cadherin domains by the LDRE and D*X*ND motifs of one domain and the D*X*D motif of the next one. The extracellular domain of cadherin molecules appears to be stabilized by the binding of multiple calcium ions (20).

The segment of similarity to cadherins previously identified in RET extended between the V*X*Y*X*V motifs of two consecutive cadherin domains (*dotted line* in Fig. 2*B*). Visual inspection of flanking residues upstream and downstream of this segment in our multiple sequence alignment of RET ectodomains revealed many of the signature motifs of the cadherin consensus sequence as well as their respective spacing (Fig. 2*B*). This analysis delineated two consecutive cadherin-like domains (herein called CLD2 and CLD3, respectively) matching the consensus sequence of cadherin domains, extending from positions 166 to 387 in the mid-portion of the RET extracellular region (Figs. 2*B* and 3). CLD2 contains the conserved LDRE and D*X*ND motifs, which, together with the conserved D*X*D motif from CLD3, could form a calcium-binding module similar to that found in classical and protocadherin molecules.

Comparison of the multiple sequence alignment of RET ectodomains with the cadherin consensus sequence revealed the presence of two additional cadherin-like domains in the extracellular region of RET, one upstream of CLD2 from positions 28 to 156 (herein called CLD1) and another downstream of CLD3 from positions 401 to 516 (herein called CLD4) (Figs. 2*B* and 3). Highly conserved spacer sequences were found between CLD1 and CLD2 and between CLD3 and CLD4, which could not be assigned to cadherin-related proteins (*boxed* in Fig. 2*B*). Interestingly, a spacer sequence is also found between the first and second cadherin domains of DN-cadherin, a *Drosophila* neuronal adhesion receptor (21).

Within RET CLD1, hydrophobic as well as glycine and proline residues found in the cadherin consensus sequence and known to be of structural importance in the crystal structure of cadherin domains, are well conserved (Fig. 2*B*). On the other

---

[2] J. Anders, S. Kjær, P. Lönnerberg, and C. F. Ibáñez, unpublished results.

[3] To be consistent with previous literature, all RET amino acid numbering in this study starts at the initiator Met residue in the RET signal sequence. The first residue in the mature human RET polypeptide is Gly at position 28.
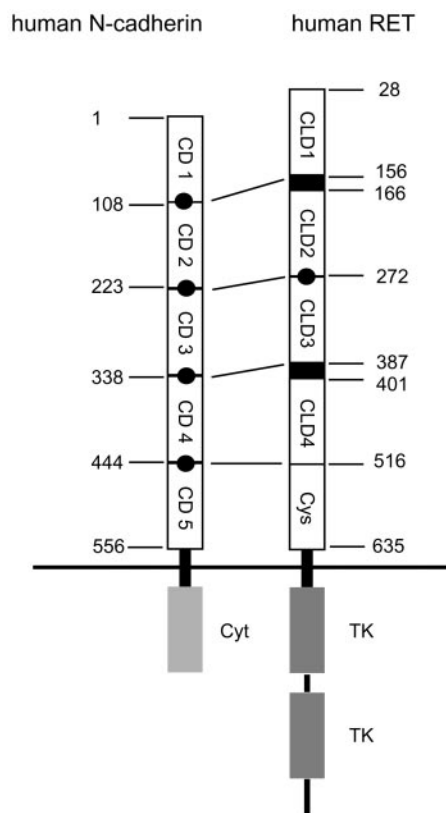
FIG. 2. **Sequence alignments of the extracellular region of cadherins and RET.** *A*, alignment of cadherin domains (*CD*) from E- and N-cadherins and protocadherin 7. At the *top* of the E- and N-cadherin alignments, the seven prototypical strands of the cadherin fold (*arrows A–G*) are indicated. Shown at the *bottom* of the *panel* is the consensus motif derived from the multiple sequence alignment. *Black shading* denotes residues identical to the consensus; *gray shading* indicates residues similar to the consensus. A long glycine-rich insertion in CD2 of protocadherin 7 between strands C and D (GGGGSGGGESRRAGAADSAPYPGGGGNGASGGGSGGSKRRLDASEGGGGTNPGGGR) is denoted by *[X]*. *B*, alignment of CLDs from vertebrate and invertebrate RET molecules. Residues identical and similar to the cadherin consensus, shown at the *bottom*, are indicated by *black* and *gray shading*, respectively. Conserved cysteines are *boxed*. Disulfide bonds found in the models of RET CLDs are indicated by *connecting lines*. Shown at the *top* of each domain are the location of the strands predicted by our molecular models. The intervening loops between CLD1 and CLD2 and between CLD3 and CLD4 of RET are *boxed*. The region of RET previously reported as having similarity to cadherins is *underlined* with a *dashed line*.

hand, the polar residues that form part of the calcium-binding site of repetitive cadherin domains are only partially conserved in CLD1 of RET (Fig. 2*B*). Together with the presence of a spacer sequence between CLD1 and CLD2 and the absence of the D*X*D motif in CLD2, this suggests that a calcium-binding site cannot be formed at the interface between CLD1 and CLD2. Compared with the other CLDs in RET and with classical cadherin domains, additional residues in the loop between strands E and F separate the LDRE and Y*X*L*XX*A*X*D motifs in CLD1 (Fig. 2*B*). A similar feature can be found in several cadherin domains of protocadherins, including the N-terminal domains of protocadherins 1 and 7 (Fig. 2*A*). Insertions within the consensus sequence seem to be a common property of a number of cadherin-related molecules. In the second cadherin domain of protocadherins 7 and 10, for example, glycine-rich sequences of up to 50 residues appear to be inserted in the loop between strands C and D, at a location where classic cadherins have a loop of 8–12 residues (8). Thus, although RET CLD1 conserves several of the features of the classical cadherin consensus sequence, it also shares intriguing properties with N-terminal domains from more divergent members of the cadherin superfamily, including several protocadherins.

In CLD4 sequences of RET from different species (Fig. 2*B*) structurally important hydrophobic, glycine, and proline residues of the cadherin consensus sequence are well conserved. Similar to CLD1, polar residues contributing to the calcium-binding sites of repetitive cadherin domains are less conserved, suggesting that CLD3 and CLD4 of RET are unlikely to form a calcium-binding site. On the other hand, unlike RET CLD1, 2 and 3, the characteristic Y*X*L*XX*A*X*D motif found in the loop between strands F and G of the cadherin consensus sequence is well conserved in vertebrate RET CLD4 (Fig. 2*B*), clearly supporting the similarity of this domain to cadherins.

Validation of RET Cadherin-like Domains by Data Base Searches—Additional support for the existence of four cadherin-like domains in the extracellular region of RET was obtained from data base searches using different algorithms. A search using the sequences of human RET cadherin-like domains in the GenPept Data Base with the PSI-BLAST algorithm returned significant alignments to members of the cadherin superfamily. Cadherin consensus sequences and domain borders were perfectly aligned between human RET cadherin-like domains and cadherins. Importantly, no other sequences were recovered with a higher score than either RET itself or

human N-cadherin   human RET

TABLE I

Statistical quality factors of modeled cadherin-like domain structures of the extracellular region of RET

Modeled structure	Structural average quality control value[a]	Average G factor[b]	Total free energy[c]
			kJ/mol
CLD1	-1.28	-0.49	-6555
CLD2 + CLD3	-1.52	-0.47	-12328
CLD4	-1.515	-0.39	-4775

[a] As calculated by WHATIF (29). The structural average quality control factor represents packing quality of the residues and is reported to be in the range of -2 to -1 for modeled structures.

[b] As calculated by PROCHECK (30). The G factor is representing the stereochemical parameters of a given structure including torsion angles, main chain bond lengths, and main chain bond angles. The values should be greater than -0.5.

[c] As calculated using the GROMOS 96 implementation in the Swiss-PdbViewer program suite (15). The total free energy of a modeled structure should be negative in case of thermodynamically allowed structures.

cadherin-related molecules. The sequence recovered with the most significant score was mouse cadherin-related neuronal receptor 1, a member of a cadherin subfamily of ~20 genes expressed at neuronal synapses (22). Cadherin domains 3, 4, and 5 of mouse cadherin-related neuronal receptor 1 (Gen-Bank[TM] accession code CD86916; $E = 7 \times 10^{-113}$) aligned with 18% amino acid identity to human RET CLD1, CLD2, and CLD3, and the spacer sequence separating CLD1 and CLD2 of RET was inserted at the predicted position. Zebrafish RET was detected with a comparable score, $E = 4 \times 10^{-124}$.

Structural data base searches and fold predictions also confirmed the identification of four cadherin-like domains in the extracellular region of RET. Using the collection of vertebrate RET cadherin-like domains to search the Structural Classification of Proteins Data Base, we found significant assignments to cadherin domains for all vertebrate RET molecules. Using the three-dimensional position-specific scoring matrix server to predict the fold of RET cadherin-like domains by comparison with the Brookhaven protein structure data base (14), all vertebrate RET domain sequences returned significant annotations to the cadherin fold for all four RET cadherin-like domains. Neither search returned assignments to any other protein family with higher statistical significance than to cadherin domains.

Comparison of the Domain Organization of RET Extracellular Domains and Classical Cadherins—The domain organization proposed here for RET has a striking resemblance to that of the classical neural cadherin (N-cadherin). Instead of a cys-

teine-rich domain, N-cadherin has a more divergent fifth cadherin domain close to the membrane (Fig. 3). Domain borders are found at almost identical distances from the plasma membrane in both RET and N-cadherin, with the small differences being accounted for the unique spacer sequences found in RET between CLD1 and CLD2 and between CLD3 and CLD4 (Fig. 3). Although the cadherin domains in N-cadherin are strictly repetitive, allowing for the formation of four calcium-binding sites, RET has only one conserved calcium-binding motif between CLD2 and CLD3.

Based on the crystal structures of epithelial and neural cadherins, the N-terminal domain of classical cadherin molecules has been proposed to mediate cell adhesion via homophilic interactions involving the conserved recognition sequence HAV in the C-terminal region of the domain (part of the consensus YXLXXAXD motif in strand F; see Fig. 32 2*A*) (23). However, the HAV sequence is not generally conserved among members of the cadherin superfamily (see *Protocadherin 7* in Fig. 2*A*) and is also absent in the proposed cadherin-like domains of RET. The possibility that RET may be involved in homophilic interactions mediating cell adhesion remains to be tested.

Molecular Modeling of RET Cadherin-like Domains—We used the crystal structures of domains 1 and 2 from epithelial and neural cadherins as templates to model the structures of the four cadherin-like domains identified in the extracellular region of RET. Structural alignments of the cadherin-like domains of human RET to the template sequences were generated by superimposing the identified consensus residues of cadherin domains in both sets of sequences. The sequence identity ranged between 16 and 23%, and the sequence similarity was between 42 and 50%. Although RET is only a distant member of the cadherin superfamily, clusters of hydrophobic, polar, and charged residues showed a remarkable degree of overlap in the alignment, clearly supporting a similar fold. Moreover, proline and glycine residues, which indicate an interruption in the secondary structure, were also often superimposed in the structural alignment. Models were constructed as described under "Materials and Methods." The statistical quality factors of the models are shown in Table I.

The model of RET CLD2 and CLD3, containing a putative calcium-binding site, was found to deviate with a root mean square Cα coordinate difference of 2 Å from domains 1 and 2 of the N-cadherin template structure. RET CLD1 was more similar to domain 2 of classical cadherins, and its modeled structure deviated with a root mean square Cα coordinate difference of 2.2 Å from the second domain of N-cadherin. On the other hand, RET CLD4 showed higher similarity to the N-terminal domain of classical cadherins, and its modeled structure devi-
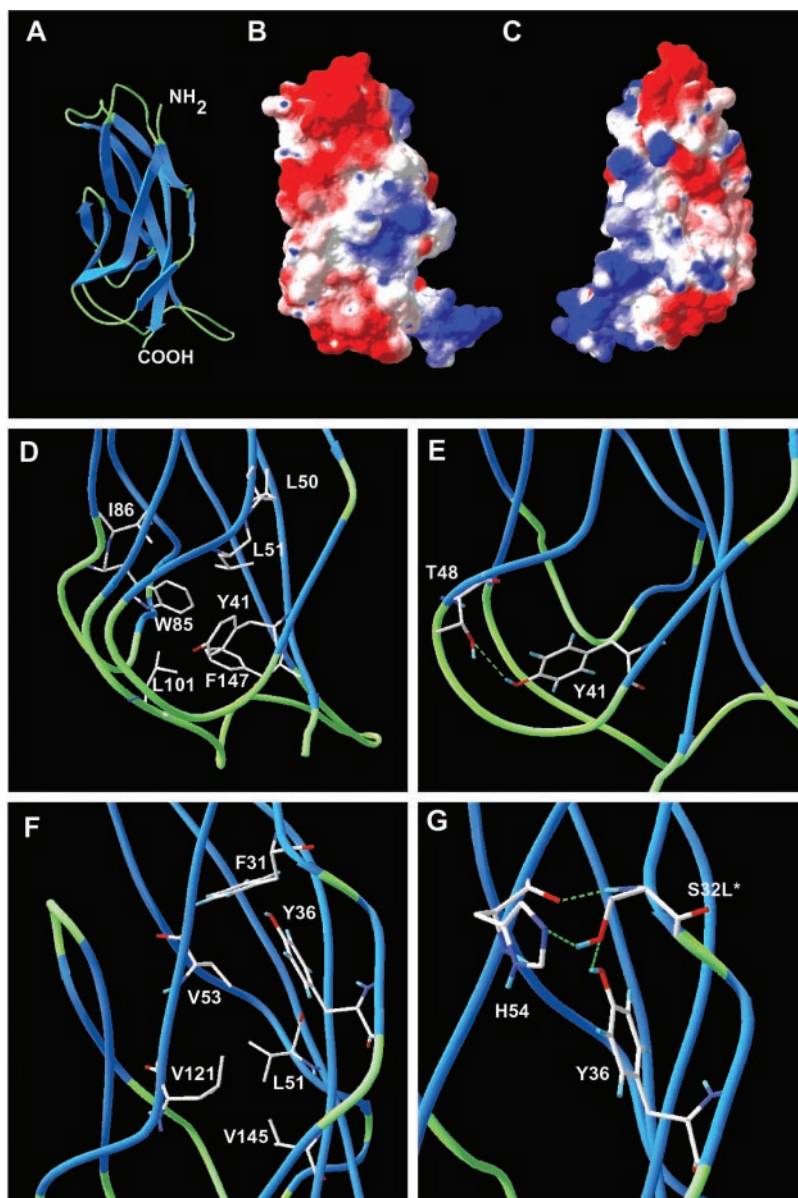
FIG. 4. **Modeled structure of human RET CLD1.** *A*, ribbon representation of the model of the first CLD of human RET from amino acids 28 to 156. Loop structures are represented as *green tubes*. β-Strands are shown in *blue*. *B*, Coulomb electrostatic potential plotted on the solvent-accessible surface of the modeled structure in the same orientation as the ribbon presentation shown in *A*. *C*, same as *B* but rotated 180°. *D–G*, enlarged views of hydrophobic cores and hydrogen bond networks of conserved tyrosine residues in the model (see text for details). *Asterisks* denote residues mutated in HSCR.

ated 2.8 Å from domain 1 of the N-cadherin template. We present below a detailed description of the molecular models of individual RET cadherin-like domains.

Modeled Structure of RET CLD1—The model of the first cadherin-like domain of human RET is shown in Fig. 4. The insertion downstream of the LDRE consensus motif (which in human RET CLD1 is LDHS; Fig. 2*B*) forms an extended loop protruding from the side of the domain (Fig. 4*A*). Most hydrophobic residues that are either conserved or replaced by similar residues in vertebrate RET molecules are buried in the interior of the model of RET CLD1. Four conserved hydrophobic residues, Tyr[30], Leu[40], Val[42], and Tyr[96], are solvent-exposed. The completely conserved residues Leu[95] and Leu[97], flanking the solvent exposed Tyr[96], are buried in the model. Leu[95] forms part of the cadherin consensus motif TGXL (strand E) and is also buried in the crystal structures of epithelial and neural cadherins (9–12). The conserved Tyr[41], flanked by solvent-exposed Leu[40] and Leu[42], is buried in the model and forms part of a hydrophobic core at the bottom of the domain, together with other conserved hydrophobic residues, including Leu[50], Leu[51], Trp[85], Ile[86], Leu[101], and Phe[147] (Fig. 4*D*). Leu[101] forms part of the LDRE cadherin consensus motif and plays the same structural role in the hydrophobic core of domain 2 of N-cad-

herin (9). The side chain hydroxyl group of Tyr[41] is engaged in a hydrogen bond with the side chain hydroxyl group of Thr[48], thus stabilizing the modeled structure (Fig. 4*E*). A second hydrophobic core in the upper part of the CLD1 model is formed by the conserved residues Phe[31], Tyr[36], Leu[51], Val[53], Phe[66], Val[121], and Val[145] (Fig. 4*F*). In this case, structural stability is provided by a hydrogen bond between the side chain hydroxyl groups of Tyr[36] and Ser[32] (Fig. 4*G*). Finally, all conserved or conservatively replaced hydrophilic residues and all asparagine residues in putative *N*-glycosylation consensus motifs in RET CLD1 were solvent-exposed in the model.

Several mutations have been described in RET CLD1 of HSCR patients that affect the folding, maturation, and membrane transport of RET. They include S32L, L40P, P64L, R77C, and G93S; G93S was found in a sporadic case of the disease (Table II). In our model of human RET CLD1, the side chain of Ser[32] is engaged in hydrogen bonds with both the imino group of His[54] and the side chain hydroxyl group of Tyr[36], probably contributing to the stabilization of the domain (Fig. 4*G*). Intriguingly, Ser[32] is only conserved among mammalian RET variants (Fig. 2*B*); it is replaced by proline in chicken, zebrafish, and *Drosophila* and by leucine, *i.e.* the same residue found in Hirschsprung's disease patients, in *Xenopus* RET

TABLE II

Mutations found in HSCR patients within cadherin-like domains of the extracellular region of RET and the structural context of the residues as predicted by our models

HSCR mutation	Localization	Shown *in vitro* to impair RET folding and maturation	Reference	Conserved in different species	Conserved in cadherins	Predicted structural context
S32L	CLD1	Yes	. 5	No	No	Hydrogen bond
L40P	CLD1	Yes	. 31	No (V, I)	No	β-Sheet
P64L	CLD1	Yes	. 31	Yes	No	Constraint in loop
R77C	CLD1	Yes	. 32	No	No	False disulfide bond
G93S	CLD1	No	. 33	Yes	Yes	Constraint in loop
C197Y	CLD2	No	. 25	Yes	No	Disulfide bond
R231H	CLD2	Yes	. 34	Yes	Yes	Hydrogen bond
D264K	CLD2	Yes	. 34	Yes	Yes	Calcium binding
R287K	CLD3	Yes	. 34	Yes	No	Hydrogen bond
D300K	CLD3	Yes	. 34	Yes	Yes	Calcium binding
R330Q	CLD3	Yes	. 31	In tetrapods	No	Hydrogen bond
R360W	CLD3	Yes	. 32	No	No	Hydrogen bond
P399L	CLD4	No	. 35	Yes	Yes	Constraint in loop
D469N	CLD4	No	. 36	Yes	No	Loss of negative charge
R475Q	CLD4	No	. 37	Yes	Yes	Hydrogen bond

(see alignment in Fig. 2*B*). Thus, other changes may compensate for these replacements in the RET molecules of these species. *In silico* replacement of Leu[40] with proline in the model of RET CLD1 caused the disruption of the β-strand (third segment of strand A in Fig. 2*B*) in which Leu[40] is located and prevented the formation of the hydrophobic core around Tyr[41] (Fig. 4*D*). Pro[64] is located in a long loop containing the conserved Phe[66], which forms part of the hydrophobic core of the domain, and our model indicates that the P64L HSCR mutation could cause displacement of Phe[66] away from the hydrophobic core, thereby affecting the stability of the domain. Regarding the R77C mutation, the additional cysteine would be in a position to disrupt the model structure by forming an additional disulfide bond with nearby cysteines. Finally, *in silico* replacement of Gly[93], part of the TG*X*L cadherin consensus motif in strand E, with Ser, to mimic the G93S mutation found in a sporadic HSCR case, resulted in clashes of neighboring side chains that could not be removed by rotations around side chain torsion angles. In addition, this mutation buried an unsatisfied hydrogen bonding donor/acceptor (*i.e.* the serine side chain) that is very unfavorable. Thus, our model predicts that all five HSCR mutations in CLD1 affect structurally important residues and would destabilize the domain.

Analysis of Selected Residues in RET CLD1 by Site-directed Mutagenesis—We tested experimentally some of the predictions of our model of RET CLD1 using site-directed mutagenesis. Tyr[36], Tyr[41], and Trp[85] are predicted to be structurally critical residues because of both their participation in the hydrophobic core of the domain and their ability to form hydrogen bonds that stabilize RET CLD1. Tyr[30], on the other hand, is one of the few conserved hydrophobic residues that our model predicts to be solvent-accessible and therefore unlikely to play a structural role in the stability of the domain. Tyr[36] was mutated to serine, whereas Tyr[30], Tyr[41], and Trp[85] were mutated to alanine, and each mutation was tested for its role in RET folding and maturation. This was facilitated by the observation that mutations that disrupt RET folding invariably result in the accumulation of a lower molecular mass partially glycosylated species that does not reach the membrane but is instead retained in the endoplasmic reticulum (5, 24). In addition to its higher mobility, this species is also sensitive to Endo H deglycosylase, an enzyme specific for the glycosylation that occurs before the exit of proteins from the endoplasmic reticulum (24). When ectopically overexpressed in mammalian cells, the extracellular domain of wild type RET yielded both the fully matured 130-kDa protein and variable amounts of the partially processed 105-kDa species (Fig. 5*A*). On the other hand, the
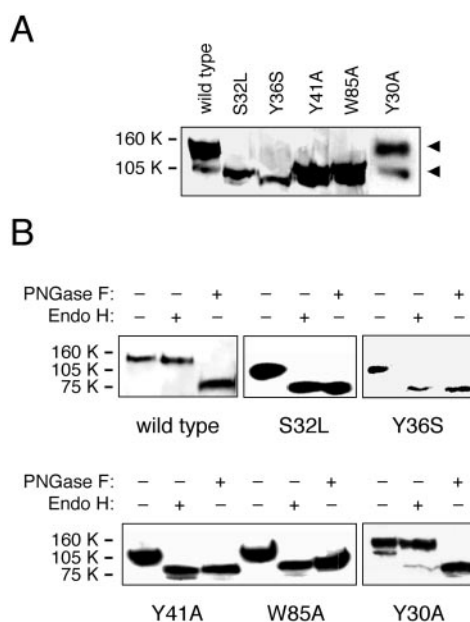

FIG. 5. **Functional analysis of hydrophobic residues in RET CLD1 by site-directed mutagenesis.** *A*, misfolded RET extracellular domain mutants migrate faster than wild type because of incomplete glycosylation. Fifty microliters of a 10-fold concentrated conditioned medium from Chinese hamster ovary or HEK cells stably expressing the wild type RET extracellular domain or the indicated mutants was loaded, blotted, and subsequently probed with an anti-hemagglutinin antibody. A lower molecular mass is observed for some of the mutants because of incomplete glycosylation. Although normally retained in the endoplasmic reticulum, butyrate-induced overexpression and cell lysis contributed to the release of this species to the culture supernatant. *B*, misfolded RET extracellular domain mutants are sensitive to treatment with Endo H deglycosylase. Fifty microliter of conditioned medium containing wild type or mutant RET extracellular domain was subjected to a deglycosylation treatment using either Endo H or Peptide:N-glycosidase F as indicated. Only wild type RET was resistant to the Endo H treatment, confirming that the mutants accumulate as glycosylation intermediates.

Hirschsprung's disease mutant S32L, which disrupts folding and maturation of the RET protein (5), yielded primarily the lower molecular mass species (Fig. 5*A*). Moreover, the S32L mutant, but not the wild type RET extracellular domain, was sensitive to deglycosylation by Endo H (Fig. 5*B*). Both proteins were sensitive to deglycosylation by Peptide:N-glycosidase F, which deglycosylates all sugars regardless of maturation state and reduces the apparent molecular mass of the extracellular
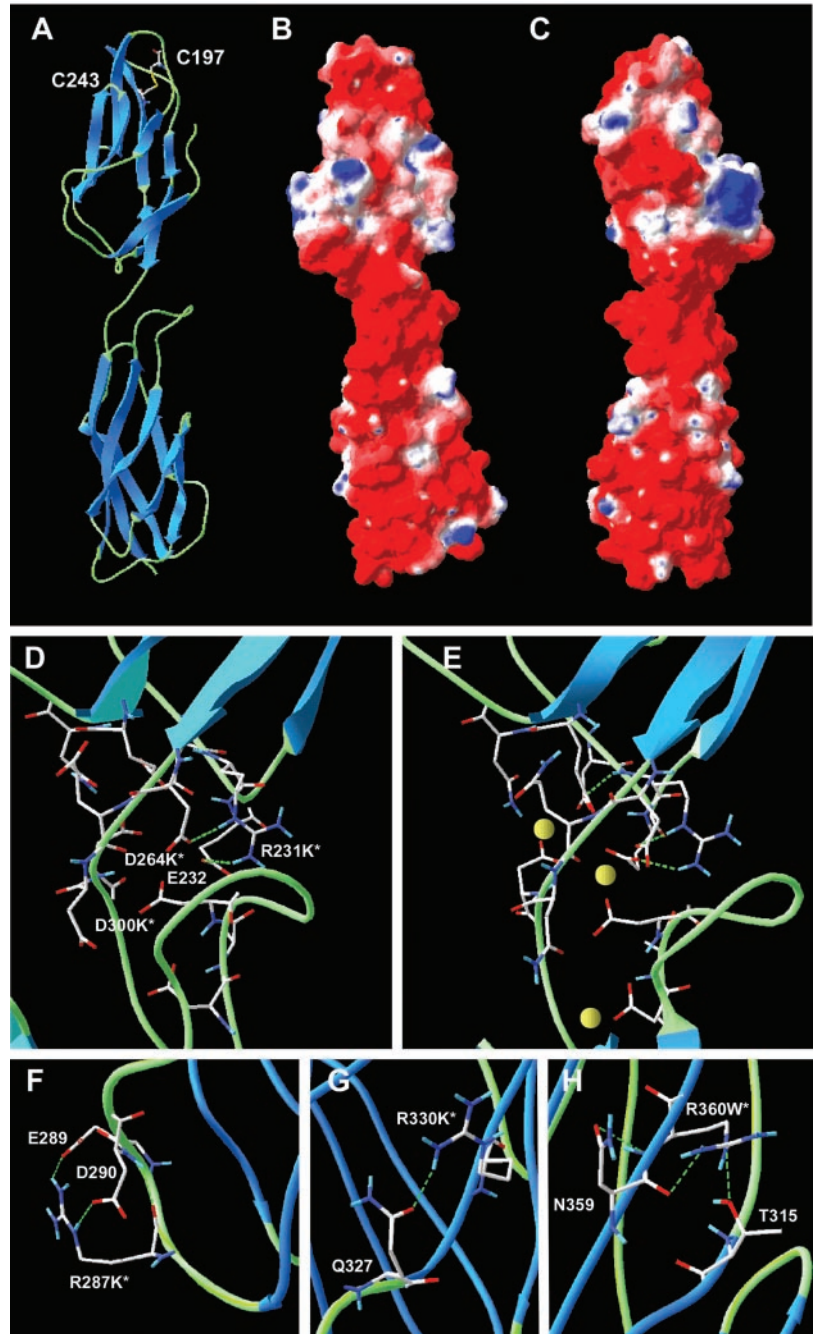
Fig. 6. **Modeled structure of human RET CLD2 and CLD3, and comparison of the calcium binding sites of RET and E-cadherin.** *A*, ribbon representation of the model of amino acids 168–386 of human RET comprising the CLD2 and CLD3. The disulfide bridge predicted by the model is shown by CPK color-coded stick-and-ball models. Loop structures are represented as *green tubes*. β-Strands are shown in *blue*. *B*, Coulomb electrostatic potential plotted on the solvent-accessible surface of the modeled structure in the same orientation as the ribbon presentation shown in *A*. *C*, same as *B* but rotated 180°. *D*, enlarged view of the calcium-binding site of RET. *Asterisks* denote residues mutated in HSCR. *E*, enlarged view of the calcium binding site of E-cadherin (12). Bound calcium ions are depicted by *yellow balls*. *F–H*, enlarged views of the hydrogen bond and salt bridge network of arginine residues mutated in HSCR, which are denoted with *asterisks* (see the text for details).

domain to 80 kDa (Fig. 5*B*). As predicted by our model, the Y36S, Y41A, and W85A mutations resulted in the accumulation of partially processed RET at expense of the fully matured species (Fig. 5*A*). In addition, the three mutations gave RET extracellular domains that were sensitive to Endo H (Fig. 5*B*). In contrast, the Y30A mutation did not affect the production of high molecular mass matured RET (Fig. 5*A*), in agreement with it being a solvent-exposed residue. These results are in agreement with our model and suggest that it will also be useful for the identification of functionally important residues involved in ligand binding and receptor activation.

Modeled Structure of RET CLD2 and CLD3—A model of the tandem CLD2 and CLD3, encompassing the putative calcium-binding site is shown in Fig. 6. As in the model of CLD1, conserved hydrophobic residues in CLD2 and CLD3 are buried, and those forming part of the cadherin consensus sequence match in role and position the corresponding residues in the crystal structures of epithelial and neural cadherins, closely

resembling the hydrophobic core of classical cadherins. The putative calcium-binding motif in RET is shown in close-up view side by side with the calcium-binding region of epithelial cadherin in Fig. 6 (*D* and *E*, respectively). All asparagine residues in putative *N*-glycosylation consensus motifs were solvent-exposed in our models of RET CLD2 and CLD3.

HSCR mutations in RET CLD2 and CLD3 for which there is biochemical evidence indicating that they disrupt the folding and maturation of RET include R231H, D264K, R287K, D300K, R330Q, and R360W (Table II). Arg[231] is the conserved arginine in the LDRE cadherin consensus motif in RET CLD2. Both in our model and in the crystal structure of epithelial cadherin (12), this arginine residue appeared engaged in salt bridges with acidic side chains of conserved aspartic and glutamic acid residues that form part of the Ca^{2+}-binding site (Fig. 6, *D* and *E*), suggesting a likely role in the stabilization of this motif in both RET and classical cadherins. Asp[264] and Asp[300] are both invariably conserved in all RET molecules and corre-

spond to aspartate residues involved in Ca^{2+} coordination in classical cadherins (Fig. 6, *D* and *E*). Mutation of either of these two residues to lysine effectively inverts the charge at these positions, thereby preventing Ca^{2+} coordination. The three arginine residues in RET CLD3 affected in HSCR mutations were found to be engaged in a network of hydrogen bonds and salt bridges with other polar and charged residues on the surface of the domain (Fig. 6, *F–H*), stabilizing the conformation of the domain. These arrangements are highly conserved in tetrapod RET sequences (with Gln[327] being replaced by Glu in chicken and *Xenopus*), supporting their structural importance.

In addition, human RET CLD2 contains three cysteine residues, two of which (Cys[197] and Cys[243]) are conserved in all RET sequences cloned so far including *Drosophila* RET. These two cysteine side chains form a disulfide bridge in our model (Fig. 6A). Intriguingly, a C197Y mutation has been found associated with a sporadic form of Hirschsprung's disease (25). Although this mutation has not been analyzed at the biochemical level, it is likely to affect the stability of the RET protein by disrupting this cysteine bridge in CLD2. Our model of RET CLD2 and CLD3 therefore predicts that all known Hirschsprung's mutations disrupt the folding of the domain and decrease its stability.

The Extracellular Domain of RET Ligates Ca^{2+}, and Ca^{2+} Binding Is Required for the Interaction of RET with GDNF—Calcium ions are important for the transport of fully matured RET protein to the plasma membrane (26) and for RET activation by GDNF (27). However, binding of Ca^{2+} ions to the extracellular domain of RET has never been directly demonstrated. Using radioactive $^{45}Ca^{2+}$ and a purified fusion protein of the extracellular domain of RET and human immunoglobulin (RET-Fc), we established that RET specifically binds Ca^{2+} (Fig. 7A). In contrast, no binding of GFRα1-Fc to Ca^{2+} could be detected above control (Fig. 7A). Ca^{2+} binding may function to stabilize the extracellular domain of RET in a conformation that facilitates ligand binding. Using chemical cross-linking of ^{125}I-GDNF to monolayers of cells expressing RET and GFRα1, we found that RET required Ca^{2+} for binding to ^{125}I-GDNF (Fig. 7B). In contrast, GDNF was still able to bind to the GFRα1 co-receptor in the absence of extracellular Ca^{2+} (Fig. 7B). Together, these observations indicate that RET binds Ca^{2+} and that this is required for the stabilization of a conformation of its extracellular domain that is required for ligand binding.

Modeled Structure of RET CLD4—The model of RET CLD4 is shown in Fig. 8. Most conserved hydrophobic residues are buried and involved in hydrophobic cores closely resembling those found in the crystal structures of classical cadherins. All asparagine residues in putative *N*-glycosylation consensus motifs are solvent-exposed. No congenital HSCR mutations have been found in CLD4 of human RET. However, three point mutations, *i.e.* P399L, D469N, and R475Q (Table II), have been found in patients with sporadic forms of the disease but have not yet been validated biochemically and will be discussed below. Pro[399] is highly conserved in RET and classical cadherins. Both in our model and in the crystal structures of E- and N-cadherins, P399 is located in a coil structure at the
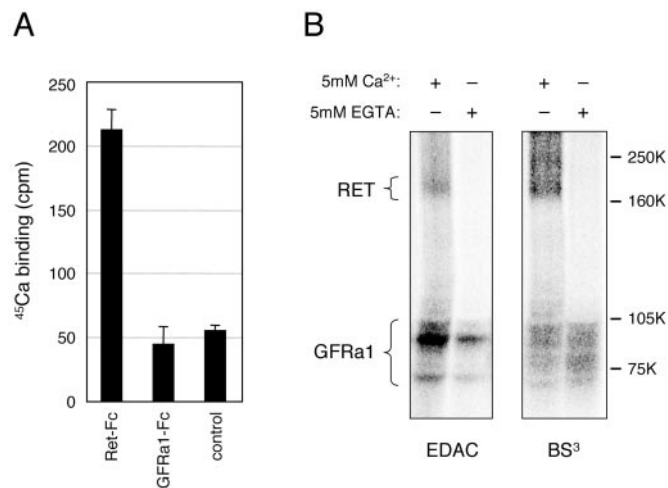

FIG. 7. **The extracellular domain of RET ligates Ca^{2+} and Ca^{2+} binding is required for the interaction of RET with GDNF.** *A*, binding of $^{45}Ca^{2+}$ to the extracellular domain of RET. $^{45}Ca^{2+}$ was added to microtiter plate wells coated with either RET-Fc, GFRα1-Fc, or only capturing antibodies (*control*). The amount of calcium retained in the wells after repeated washes was then measured. The values are the average values ± S.D. of triplicate determinations. Similar results were obtained following a prior Ca^{2+} depletion with EGTA (not shown). *B*, binding of GDNF to RET requires Ca^{2+}. Affinity labeling of endogenous RET and GFRα1 receptors on the motorneuron cell line MN1. ^{125}I-GDNF was allowed to equilibrate on MN1 cell monolayers in the presence or the absence of Ca^{2+} or the Ca^{2+} chelator EGTA as indicated and subsequently cross-linked using ethyl-dimethylaminopropyl carbodiimide (*EDAC*, *left*) or bis(succinimidyl) suberate (*BS³*, *right*) as cross-linking agents.
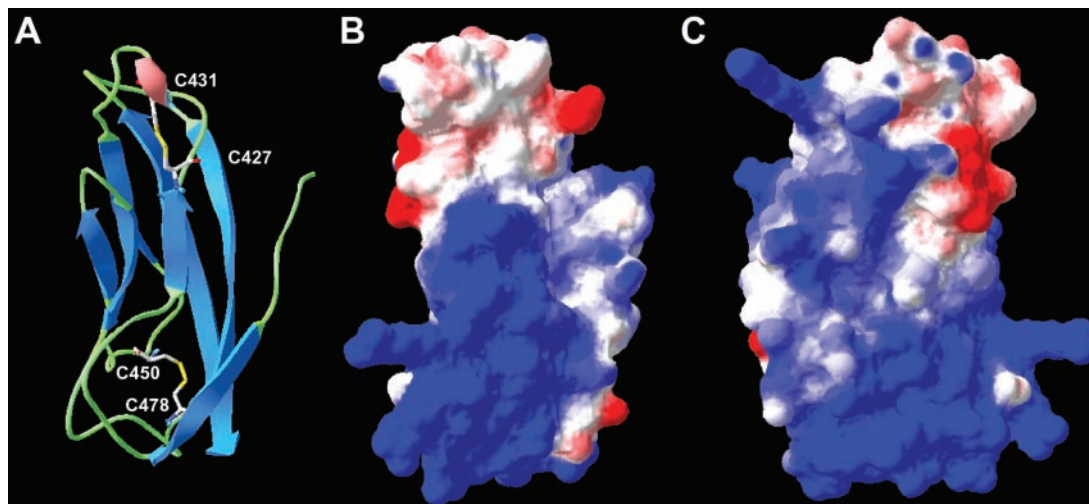

FIG. 8. **Model structure of human RET cadherin-like domain 4.** *A*, ribbon representation of the model of the fourth CLD of human RET from amino acids 401 to 516. Disulfide bridges predicted by the model are shown by CPK color-coded stick-and-ball models. Loop structures are represented as *green tubes*. β-Strands are shown in *blue*. A partial α-helix is indicated in *red*. *B*, Coulomb electrostatic potential plotted on the solvent-accessible surface of the modeled structure in the same orientation as the ribbon presentation shown in *A*. *C*, same as *B* but rotated 180°.

N terminus of the domain where it may contribute an important constraint to its fold. The side chain of Asp^{469} appears exposed to solvent and not engaged in hydrogen bonding. At present, the possible structural or functional importance of this negatively charged residue is difficult to predict, but it may stem from its privileged location within a region of otherwise positive electrostatic potential (Fig. 8, *A* and *B*). Arg^{475} forms part of the sequence L*X*R*X* that aligns with the cadherin consensus motif LDRE. A comparable HSCR mutation has also been found in RET CLD2, *i.e.* R231H (see above). Superposition of the models of RET CLD2 and CLD4 showed these two arginine residues to occupy the same position, suggesting a role for Arg^{475} in the stabilization of the domain by hydrogen bonding.

In addition, RET CLD4 contains four cysteine residues that in our model form two disulfide bonds that probably stabilize the fold of the domain (Fig. 8*A*). A short helical segment in the model of RET CLD4 allows the formation of a disulfide bond between the totally conserved Cys^{427} and Cys^{431} (Fig. 8*A*). A second disulfide bond is formed between Cys^{450} and Cys^{478}, connecting the loops between strands C and D and strands E and F (Fig. 8*A*). These two cysteine residues are conserved in all vertebrate RET sequences with the exception of *Xenopus* RET, which has the two residues replaced by serine and threonine, respectively, in agreement with their forming a disulfide bond.

Conclusions—Until now, the extracellular region of RET has been viewed as containing a short segment with similarity to cadherins amid a bulk of unrelated sequence. Using various bioinformatic tools, we have found that in fact the entire extracellular domain of the RET molecule appears to be organized in a way similar to that of cadherins and cadherin-related proteins. Our analyses have identified four repeats of a β-sandwich cadherin-like domain of about 110 residues, followed by a distinct segment, without similarity to cadherins, the so-called cysteine-rich domain. Intriguingly, the presence of cysteine-rich regions close to the plasma membrane already has a precedent in the cadherin-related receptor Flamingo from *Drosophila*, which contains nine cadherin domains followed by three membrane-proximal cysteine-rich segments (28). Thus, our model of the RET extracellular domain places this receptor as a distant, albeit genuine, member of the cadherin superfamily, not only because of its limited but significant sequence similarity to cadherins, but also, and most importantly, because of the overall architecture of its extracellular region.

RET is the only member of this superfamily containing an intrinsic protein-tyrosine kinase domain, suggesting that it may have arisen by the recombination of an ancestral cadherin with a protein-tyrosine kinase. The presence of a distant but clearly phylogenetically related RET molecule in *Drosophila* indicates that this event must have taken place during early metazoan evolution, prior to the divergence of vertebrate and invertebrate lineages. Moreover, the absence of obvious homologues of RET ligands in the *Drosophila* genome suggests that the ancestral function of this receptor might have been related to cell adhesion, as appears to be the case of most members of the cadherin superfamily.

The modeled structures presented here for CLD1, CLD2, CLD3, and CLD4 of human RET represent realistic models from both energetic and geometrical points of view and are consistent with previous observations made using RET mutants isolated from HSCR patients, as well as our own site-directed mutagenesis studies of CLD1. In the absence of exper-

imentally determined structures for the extracellular domains of RET, these models will represent valuable tools with which to guide future site-directed mutagenesis studies aimed at identifying residues involved in ligand binding and receptor activation.

REFERENCES

1. Airaksinen, M. S., Titievsky, A., and Saarma, M. (1999) *Mol. Cell. Neurosci.* **13,** 313–325
2. Saarma, M. (2000) *Eur. J. Biochem.* **267,** 6968–6971
3. Edery, P., Eng, C., Munnich, A., and Lyonnet, S. (1997) *Bioessays* **19,** 389–395
4. Santoro, W., Carlomagno, F., Melillo, R. M., Billaud, W., Vecchio, G., and Fusco, A. (1999) *J. Endocrinol. Inv.* **22,** 811–819
5. Carlomagno, F., Devita, G., Berlingieri, M. T., Defranciscis, V., Melillo, R. M., Colantuoni, V., Kraus, M. H., Difiore, P. P., Fusco, A., and Santoro, M. (1996) *EMBO J.* **15,** 2717–2725
6. Iwamoto, T., Taniguchi, M., Asai, N., Ohkusu, K., Nakashima, I., and Takahashi, M. (1993) *Oncogene* **8,** 1087–91
7. Schneider, R. (1992) *Trends Biochem. Sci* **17,** 468–469
8. Nollet, F., Kools, P., and van Roy, F. (2000) *J. Mol. Biol.* **299,** 551–572
9. Tamura, K., Shan, W. S., Hendrickson, W. A., Colman, D. R., and Shapiro, L. (1998) *Neuron* **20,** 1153–1163
10. Overduin, M., Harvey, T. S., Bagby, S., Tong, K. I., Yau, P., Takeichi, M., and Ikura, M. (1995) *Science* **267,** 386–389
11. 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) *Nature* **374,** 327–337
12. Nagar, B., Overduin, M., Ikura, M., and Rini, J. M. (1996) *Nature* **380,** 360–364
13. Page, R. D. (1996) *Comput. Appl. Biosci.* **12,** 357–358
14. Kelley, L., MacCallum, R., and Sternberg, M. (2000) *J. Mol. Biol.* **299,** 499–520
15. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723
16. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol.* **112,** 535–542
17. Trupp, M., Rydén, M., Jörnvall, H., Timmusk, T., Funakoshi, H., Arenas, E., and Ibáñez, C. F. (1995) *J. Cell Biol.* **130,** 137–148
18. Hahn, M., and Bishop, J. M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 1053–1058
19. Yoshida, K., Yoshitomo-Nakagawa, K., Seki, N., Sasaki, M., and Sugano, S. (1998) *Genomics* **49,** 458–461
20. Pertz, O., Bozic, D., Koch, A. W., Fauser, C., Brancaccio, A., and Engel, J. (1999) *EMBO J.* **18,** 1738–1747
21. Iwai, Y., Usui, T., Hirano, S., Steward, R., Takeichi, M., and Uemura, T. (1997) *Neuron* **19,** 77–89
22. Kohmura, N., Senzaki, K., Hamada, S., Kai, N., Yasuda, R., Watanabe, M., Ishii, H., Yasuda, M., Mishina, M., and Yagi, T. (1998) *Neuron* **20,** 1137–1151
23. Blaschuk, O., Sullivan, R., David, S., and Pouliot, Y. (1990) *Dev. Biol.* **139,** 227–229
24. Cosma, M. P., Cardone, M., Carlomagno, F., and Colantuoni, V. (1998) *Mol. Cell. Biol.* **18,** 3321–3329
25. Kusafuka, T., Wang, Y., and Puri, P. (1997) *J. Pediatr. Surg.* **32,** 501–504
26. van Weering, D., Moen, T. C., Braakman, I., Baas, P. D., and Bos, J. L. (1998) *J. Biol. Chem.* **273,** 12077–12081
27. Nozaki, C., Asai, N., Murakami, H., Iwashita, T., Iwata, Y., Horibe, K., Klein, R. D., Rosenthal, A., and Takahashi, M. (1998) *Oncogene* **16,** 293–299
28. Usui, T., Shima, Y., Shimada, Y., Hirano, S., Burgess, R. W., Schwarz, T. L., Takeichi, M., and Uemura, T. (1999) *Cell* **98,** 585–595
29. Vriend, G. (1990) *J. Mol. Graph.* **8,** 52–56
30. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
31. Iwashita, T., Asai, N., Murakami, H., Matsuyama, M., and Takahashi, M. (1996) *Oncogene* **12,** 481–487
32. Bolk, S., Pelet, A., Hofstra, R. M., Angrist, M., Salomon, R., Croaker, D., Buys, C. H., Lyonnet, S., and Chakravarti, A. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 268–273
33. Angrist, M., Bolk, S., Halushka, M., Lapchak, P. A., and Chakravarti, A. (1996) *Nat. Genet.* **14,** 341–344
34. Pelet, A., Geneste, O., Edery, P., Pasini, A., Chappuis, S., Attie, T., Munnich, A., Lenoir, G., Lyonnet, S., and Billaud, M. (1998) *J. Clin. Invest.* **101,** 1415–1423
35. Seri, M., Yin, L., Barone, V., Bolino, A., Celli, I., Bocciardi, R., Pasini, B., Ceccherini, I., Lerone, M., Kristoffersson, U., Larsson, L. T., Casasa, J. M., Cass, D. T., Abramowicz, M. J., Vanderwinden, J. M., Kravcenkiene, I., Baric, I., Silengo, M., Martucciello, G., and Romeo, G. (1997) *Hum. Mutat.* **9,** 243–249
36. Inoue, K., Shimotake, T., and Iwai, N. (2000) *Am. J. Med. Genet.* **93,** 278–284
37. Eng, C., and Mulligan, L. M. (1997) *Hum. Mutat.* **9,** 97–109