The CEACAM1 N-terminal Ig domain mediates cis- and trans-binding and is essential for allosteric rearrangements of CEACAM1 microclusters

Esther Klaile,1 Olga Vorontsova,1 Kristmundur Sigmundsson,2 Mario M. Müller,1 Bernhard B. Singer,3 Lars-Göran Öfverstedt,1 Stina Svensson,1 Ulf Skoglund,1 and Björn Öbrink1

1Department of Cell and Molecular Biology, Karolinska Institutet, 171 77 Stockholm, Sweden
2Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark
3Department of Anatomy, University Hospital Essen, 45147 Essen, Germany

Introduction

Tissue structure, cellular behavior, and cell function are regulated by homotypic and heterotypic intercellular interactions mediated by cell adhesion molecules (CAMs). Knowledge about CAM-mediated transmembrane signaling has medical implications because it will allow tailored design of therapeutic agents that can target specific CAMs. For a large number of CAMs, the molecular/biochemical properties are known in great detail, and crystal structures have been reported for many CAM ectodomains (Xiong et al., 2001; Boggon et al., 2002; Tan et al., 2002; Soroka et al., 2003; Xiao et al., 2004; Fedarovich et al., 2006; Korotkova et al., 2008). However, with the exception of some integrins (Kim et al., 2003; Takagi et al., 2003; Xiao et al., 2004), this has not yet given satisfactory explanations for mechanisms of ectodomain-initiated signal generation. Signaling by single-pass CAMs belonging to the immunoglobulin superfamily remains a mystery and requires additional information on the structural dynamics and supramolecular organization of native CAMs at the cell surface and how these properties are influenced by homophilic and heterophilic CAM interactions. To achieve this goal, x-ray crystallography has to be complemented by other methods that give information on individual molecules in large populations.

Members of the carcinoembryonic antigen (CEA) family, a subfamily within the immunoglobulin superfamily, play important roles in morphogenesis (Yokoyama et al., 2007),...
vasculogenesis (Gu et al., 2009), angiogenesis (Horst et al., 2006), cell proliferation (Scheffrahn et al., 2005), cell motility (Ebrahimnejad et al., 2004; Klaile et al., 2005; Müller et al., 2005), apoptosis (Kirshner et al., 2003; Singer et al., 2005), tumor growth (Leung et al., 2008), invasion (Ebrahimnejad et al., 2004), infection, and inflammation (Gray-Owen and Blumberg, 2006). The primordial molecule of the CEA family, CEA-related CAM 1 (CEACAM1), is a single-pass transmembrane type I glycoprotein, which, like many immunoglobulin-like (Ig) CAMs, is expressed as differentially spliced isoforms (Singer and Lucka, 2005; Gray-Owen and Blumberg, 2006). The two major isoforms, CEACAM1-4L and CEACAM1-4S, which differ only in their cytoplasmic domains, have ectodomains comprised of four glycosylated Ig domains. CEACAM1-induced cell signaling is regulated by its intercellular homophilic binding at the cell surface (Gray-Owen and Blumberg, 2006), which is mediated by the N-terminal Ig domain (D1) in a reciprocal D1–D1 interaction (Wikström et al., 1996; Watt et al., 2001). However, the mechanism of this adhesion-initiated signaling is still unknown.

In this study, we have approached the first step of CEACAM1 transmembrane signaling by analysis of the dynamics and kinetics of the structure and homophilic interactions of the CEACAM1 ectodomain using a combination of surface plasmon resonance (SPR)–based binding analyses, molecular electron tomography, and chemical cross-linking. We found that the CEACAM1 ectodomain is highly flexible, participating in a limited set of structurally well-defined homophilic binding interactions that give rise to two different kinds of dimers as well as trimers and higher order oligomers. When the CEACAM1 ectodomain was associated with liposomal membranes, it became organized in multimeric microclusters with a narrow size distribution. Upon CEACAM1-mediated trans-homophilic membrane adhesion, the level of parallel CEACAM1 cis-dimers increased, and the average number of molecules per cluster decreased. Together, our data provide for the first time evidence for an allostery-based mechanism for adhesion-triggered transmission of signals via reorganization of the cis-assembly of the CEACAM1 ectodomains in the plasma membrane.

**Results**

### Homophilic binding properties of CEACAM1 ectodomains characterized by SPR

The homophilic binding properties of CEACAM1 ectodomains were analyzed by SPR-based flow cell biosensor analysis. D(1–4) and D(2–4) CEACAM1 ectodomain Fc fusion proteins were immobilized as ligands on a BIAcore chip, and both His-tagged (Fig. 1) and Fc fusion ectodomains (not depicted) were used as soluble analytes. The rat D(1–4) proteins bound specifically to immobilized rat D(1–4) (Fig. 1) but not to rat D(2–4) (not depicted). No explicit binding of the rat D(2–4) constructs was observed either to rat D(1–4) or rat D(2–4) ligands (unpublished data). Thus, the recordable homophilic binding must be caused by D1–D1 interactions. The D(1–4) binding was characterized by rapid on and off rates in the presence of both EDTA and Ca/Mg, but the extent of binding was larger in Ca/Mg (Fig. 1). In addition, the divalent cations induced a more complex binding pattern, with a slower binding superimposed on the dominant rapid association/dissociation, demonstrating that at least two different binding reactions occurred. The Ca/Mg effect was not influenced by the His tag because the same divalent cation dependence was seen when CEACAM1 Fc fusion proteins, which lack a His tag, were used as analytes (unpublished data).

The association/dissociation-binding profiles can be analyzed by a global curve-fitting procedure to get information about the underlying binding reactions and their association/dissociation rate constants, provided that appropriate curve-fitting algorithms are used. Fitting to a simple 1:1 binding model (BIAevaluation 3.1) did not give satisfying results, which of course would not be expected for a situation in which the same...
reactions occur between both the analyte molecules in the soluble bulk phase and between the soluble analyte and the immobilized ligand. Therefore, we had to design novel algorithms based on a set of plausible homophilic binding reactions that reflect such a case. Three different reaction schemes were developed: (1) formation of one class of dimers by one homophilic binding site (model 1); (2) formation of two classes of dimers (types 1 and 2) by two different homophilic binding sites (model 2); and (3) formation of two classes of dimers (types 1 and 2) and one class of trimers by two different homophilic binding sites (model 3). Formation of type 1 and type 2 dimers are characterized by a set of kinetic rate constants, $k_a1/k_d1$ and $k_a2/k_d2$, respectively. Detailed reaction schemes and the derived equations are presented in Fig. S1 and Materials and methods. It should be emphasized that we apply a minimal number of parameters in these models and that the number of free-running parameters were the same in models 2 and 3.

In the presence of EDTA, all three models gave equally good curve fitting (Fig. 1 A) with no significant differences in the $x^2$ values calculated for all data points in the association and dissociation phases. Model 2 was the least plausible because it gave unacceptable large standard deviations for the rate constants. Analysis according to model 3 gave acceptable standard deviations but indicated that the formation of type 2 dimers was almost insignificant. Thus, we conclude that in the presence of EDTA, essentially only the formation of type 1 dimers was recorded. Analysis according to model 1 gave the following values of the binding constants in EDTA: $k_a1 = 0.0453 \pm 0.000033 \, \mu M^{-1} \, s^{-1}$; $k_d1 = 0.5083 \pm 0.0078 \, s^{-1}$; and $K_D1 = 11.21 \, \mu M$.

In the presence of Ca/Mg, model 3 (the trimer model) clearly gave the best fit to the experimental values with the lowest $x^2$ (Fig. 1 B), which demonstrated that both type 1 and 2 reactions were recorded under these conditions. Fitting according to the trimer model gave the following values of the binding constants in Ca/Mg: $k_a1 = 0.0890 \pm 0.0034 \, \mu M^{-1} \, s^{-1}$; $k_d1 = 0.6806 \pm 0.0074 \, s^{-1}$; $K_D1 = 7.65 \, \mu M$; $k_a2 = 0.0000598 \pm 0.0000033 \, \mu M^{-1} \, s^{-1}$; $k_d2 = 0.01241 \pm 0.00038 \, s^{-1}$; and $K_D2 = 208 \, \mu M$. Thus, these results demonstrate that CEACAM1 D(1–4) ectodomains participate in two different, simultaneously occurring homophilic binding reactions. From the kinetic rate constants, it could be determined that the equilibrium concentrations of type 1 dimers were significantly higher than those of type 2 dimers at all protein concentrations.

Structures of CEACAM1 ectodomains determined by molecular electron tomography

The structures of soluble rat CEACAM1 ectodomains containing all four Ig domains, D(1–4), or lacking the N-terminal Ig domain, D(2–4), were determined by molecular electron tomography of vitrified specimens. 3D images were reconstructed by filtered back-projection and refined by constrained maximum entropy tomography (COMET), which allowed visualization of molecular details at a resolution of $\sim 20$ Å. Particles were selected for structural analysis by two independent procedures: (1) gray-level thresholding and (2) seeded watershed segmentation (SWS; Fig. 2, A and B). The recombiant CEACAM1 D(1–4) ectodomain that was polydisperse as a result of varying degrees of glycosylation had an average molecular mass of 81 kD (Fig. S2), which corresponds to a volume of $\sim 580$ voxels (1 voxel = 5.74 x 5.74 x 5.74 Å$^3$). All CEACAM1 ectodomains should be found among particles having a volume $\geq 500$ voxels. With this constraint, gray-level thresholding and SWS returned the same particles, corresponding to CEACAM1 monomers and dimers (Fig. 2, A and B). In addition to these protein molecular particles, a large number of smaller background structures with volumes up to 500 voxels were recorded by both approaches (Fig. 2, A and C). Similar amounts of particles of the same size distribution were seen in plain buffer specimens (Fig. 2 D). Because the
yet allow recognition of the contacting interfaces even at 20-Å resolution, which gives an idea of the structural details that, at best, would be seen in glycosylated CEACAM1 dimers/oligomers by electron tomography.

Individual Ig domains were resolved in many of the tomographically determined CEACAM1 ectodomains (Fig. 4 and Video 1). Because the rat CEACAM1 D1 domain is larger and has more glycosylation sites (105 amino acids and three N-glycosylation sites) than the D4 domain (67 amino acids and two N-glycosylation sites; Edlund et al., 1993), it might be possible to distinguish the two ends in individual molecules. To test this possibility, 20 different, extended monomeric D(1–4) domains were divided into four equally long segments, and the mean of the two diameters of the ellipsoidal center cross sections of each segment was recorded. This showed that one end had a significantly larger mean diameter than the other end in all molecules. Setting the mean of the larger end mean diameters to 1 gave a mean of the smaller end mean diameters of 0.81 ± 0.09 (P = 4 × 10−11). Similar analysis of D(2–4), which was divided into three segments, showed no significant difference in the size of the two end domains, the smaller having a relative size of 0.96 ± 0.05 (P = 0.09; D2 has 74 amino acids and six N-glycosylation sites). This indicates that the D1 and D4 ends can be discriminated in a large proportion of the recorded molecules.

A striking feature was the structural variability of the ectodomains, which was caused by a large flexibility between the covalently linked Ig domains. Many shapes were observed, including extended (Fig. 4, A and B), kinked (Fig. 4 C), and completely back-folded molecules (Fig. 4 D). All molecules in several reconstructed volumes were analyzed, which allowed quantification of the different shape categories (Fig. 4 L). In both Ca/Mg and EDTA environments, monomeric D(1–4) was buffer controls showed no particles of or above the size of CEACAM1 ectodomains, all CEACAM1 molecules could unambiguously be identified and analyzed.

To demonstrate which molecular details that ideally would be observed at a resolution of 20 Å, we show published crystallographic structures of mouse CEACAM1 (Tan et al., 2002), human CEACAM1 (Fedarovich et al., 2006), and human CEACAM5 (Fig. 3, A–C; Korotkova et al., 2008). The D1 domains of CEACAM1 and CEACAM5 are very similar both in their primary and 3D structures (Watt et al., 2001). All structures are shown both at atomic resolution (space-filling models), and at 20-Å resolution (nuclear scattering models). The two tandem domains of mouse CEACAM1 (Fig. 3 A) can easily be distinguished at 20-Å resolution. However, it should be noted that the displayed crystallographic structures represent non-glycosylated protein domains, whereas the molecular tomography of rat CEACAM1 in this study was made on highly glycosylated proteins. The tomographic structures are therefore expected to be somewhat larger and not as well resolved. Ig domains have two faces consisting of closely opposed sheets of β strands, CFG, and ABED faces (Fig. 3 D). Fig. 3 B shows two unglycosylated CEACAM5 D1 domains in close contact at their CFG faces. It is believed that this represents a physiological interaction because mutagenesis experiments have demonstrated that amino acid residues at the CFG surface are involved in homophilic adhesion (Watt et al., 2001). Fig. 3 C shows two unglycosylated CEACAM1 D1 domains that are in contact via their ABED faces. It has been argued that this does not represent a physiological interaction because the ABED surface is highly hydrophobic and is believed to be covered by an oligosaccharide attached to Asn70 (Tan et al., 2002; Fedarovich et al., 2006). Both of the structures (Fig. 3, B and C) are compact yet allow recognition of the contacting interfaces even at 20-Å resolution, which gives an idea of the structural details that, at best, would be seen in glycosylated CEACAM1 dimers/oligomers by electron tomography.

Individual Ig domains were resolved in many of the tomographically determined CEACAM1 ectodomains (Fig. 4 and Video 1). Because the rat CEACAM1 D1 domain is larger and has more glycosylation sites (105 amino acids and three N-glycosylation sites) than the D4 domain (67 amino acids and two N-glycosylation sites; Edlund et al., 1993), it might be possible to distinguish the two ends in individual molecules. To test this possibility, 20 different, extended monomeric D(1–4) domains were divided into four equally long segments, and the mean of the two diameters of the ellipsoidal center cross sections of each segment was recorded. This showed that one end had a significantly larger mean diameter than the other end in all molecules. Setting the mean of the larger end mean diameters to 1 gave a mean of the smaller end mean diameters of 0.81 ± 0.09 (P = 4 × 10−11). Similar analysis of D(2–4), which was divided into three segments, showed no significant difference in the size of the two end domains, the smaller having a relative size of 0.96 ± 0.05 (P = 0.09; D2 has 74 amino acids and six N-glycosylation sites). This indicates that the D1 and D4 ends can be discriminated in a large proportion of the recorded molecules.

A striking feature was the structural variability of the ectodomains, which was caused by a large flexibility between the covalently linked Ig domains. Many shapes were observed, including extended (Fig. 4, A and B), kinked (Fig. 4 C), and completely back-folded molecules (Fig. 4 D). All molecules in several reconstructed volumes were analyzed, which allowed quantification of the different shape categories (Fig. 4 L). In both Ca/Mg and EDTA environments, monomeric D(1–4) was
the dominant species (45.2% and 56.1% of the molecules, respectively). In addition to the monomers, several complexes of homophilically bound D(1–4) ectodomains occurred, as judged both from the shapes and dimensions/volumes of the recorded particles. The most common complex type was fully or partly extended dimers held together only by reciprocal binding between the largest end domains (Fig. 4 E). They were referred to as C-dimers because their structure and shape can easily account for antiparallel intermembrane adhesion, which is mediated by mutual D1–D1 binding between the CFG β-sheet faces (Fig. 3 B; Watt et al., 2001). Another type of D(1–4) dimers, referred to as A-dimers, was characterized by in-register, close binding between three or all four of the Ig domains (Fig. 4, F and G). Because the liposome adhesion experiments (see...
and EDTA, respectively. This shows that both types of dimers can form in the presence and absence of divalent cations, but that both dimerization reactions were enhanced by the divalent cations, which agrees with the SPR data. Furthermore, the homophilic binding affinity leading to C-dimer formation was significantly higher than that for A-dimer formation. This suggests that the C-dimers and A-dimers correspond to the type 1 and type 2 dimers recorded in the SPR-binding experiments, respectively, because the equilibrium concentrations of type 1 dimers were significantly higher than those of type 2 dimers.

The equilibrium-binding constants obtained by SPR and molecular tomography were in very good agreement. In fact, the values were surprisingly similar given the vastly different principles that the two methods are based on. This strongly supports the interpretations of the structures observed by molecular tomography.

Adhesion mediated by liposome-anchored CEACAM1

To gain further insights into the mechanism of CEACAM1-mediated adhesion, we investigated His-tagged rat CEACAM1 ectodomains anchored to nickel–nitrilotriacetic acid (Ni-NTA) liposomes. D(1–4), but not D(2–4), induced adhesion between liposome vesicles, as demonstrated both by turbidity measurements (Fig. 5A) and 2D electron microscopy (Fig. 5B). The inability of D(2–4) to mediate adhesion confirms previously published data (Wikström et al., 1996) that domain D1 mediates trans-homophilic binding and that antiparallel binding involving domains D2, D3, and D4 does not occur. The rate and extent of D(1–4)-mediated liposome aggregation increased with increasing protein/lipid ratio (1:5). The small black dots indicate 10-nm colloidal gold particles. Bars, 100 nm.

More than 90% of the truncated D(2–4) ectodomains appeared as extended (Fig. 4I) or more condensed, kinked monomers (Fig. 4J). Some dimers occurred (Fig. 4K) but no trimers. In all observed dimers, the two ectodomains were oriented in a parallel manner, suggesting that they corresponded to the A-dimers formed by D(1–4) (Fig. 4G). The lack of trimers gives further support to the notion that the trimeric structures occurring in the D(1–4) specimens were formed by reciprocal D1-binding interactions. Collectively, these results demonstrate that all four Ig domains can participate in parallel, in-register cis-dimerization, whereas antiparallel trans-binding is mediated by the D1 domain alone.

Because the input concentration of the CEACAM1 D(1–4) ectodomain was known, the molar concentrations of monomers, C-dimers, A-dimers, and trimers could be calculated from their relative abundance (Fig. 4L). Accordingly, it was possible to determine the equilibrium dissociation constants for C- and A-dimer formation by applying the law of mass action. The $K_D$ for the monomer/C-dimer binding was 43.8 µM and 72.1 µM in Ca/Mg and EDTA, respectively. The $K_D$ for the monomer/A-dimer binding was 113 µM and 260 µM in Ca/Mg and EDTA, respectively. This shows that both types of dimers can form in the presence and absence of divalent cations, but that both dimerization reactions were enhanced by the divalent cations, which agrees with the SPR data. Furthermore, the homophilic binding affinity leading to C-dimer formation was significantly higher than that for A-dimer formation. This suggests that the C-dimers and A-dimers correspond to the type 1 and type 2 dimers recorded in the SPR-binding experiments, respectively, because the equilibrium concentrations of type 1 dimers were significantly higher than those of type 2 dimers.

The equilibrium-binding constants obtained by SPR and molecular tomography were in very good agreement. In fact, the values were surprisingly similar given the vastly different principles that the two methods are based on. This strongly supports the interpretations of the structures observed by molecular tomography.

Adhesion mediated by liposome-anchored CEACAM1

To gain further insights into the mechanism of CEACAM1-mediated adhesion, we investigated His-tagged rat CEACAM1 ectodomains anchored to nickel–nitrilotriacetic acid (Ni-NTA) liposomes. D(1–4), but not D(2–4), induced adhesion between liposome vesicles, as demonstrated both by turbidity measurements (Fig. 5A) and 2D electron microscopy (Fig. 5B). The inability of D(2–4) to mediate adhesion confirms previously published data (Wikström et al., 1996) that domain D1 mediates trans-homophilic binding and that antiparallel binding involving domains D2, D3, and D4 does not occur. The rate and extent of D(1–4)-mediated liposome aggregation increased with increasing protein/liposome ratio (Fig. 5A). Electron tomography revealed the presence of CEACAM1 ectodomains both on free-liposome surfaces and in contact areas between adjacent vesicles (Fig. 6). D(1–4) ectodomains on free-liposome surfaces occurred as a mixture of monomers, dimers, and clusters containing 3–12
Figure 6. Molecular tomography of CEACAM1 ectodomains attached to liposomes. Segmentation was performed by gray-level thresholding. (A) Tomogram of D[1–4]-His–decorated liposomes. The z direction is parallel to the electron beam. (B and C) Free surfaces of D[1–4]-His–decorated (B) and D[2–4]-His–decorated (C) liposomes. (D and E) Monomeric bridges (C-dimers) connecting two liposomes (L1 and L2) are shown at two angles. In the colored panels, D1–D1 contact zones are indicated by a black dotted line, and individual Ig domains are numbered 1–4. Arrowheads point to bridge positions, and arrows point to the binding areas. (F and G) Multimeric bridges connecting two liposomes (L1 and L2) are shown. The bridge in F is composed of seven molecules anchored to L1 (orange) and L2 (green). (G) Two bridging clusters of 6 and ~10 molecules. (H and I) D[1–4] (n = 130) and D[2–4] (n = 163) monomers/clusters on free-liposome surfaces and D[1–4] monomers/clusters (n = 63) bridging two membranes were analyzed and classified according to size. The differences in each cluster class compared with D[1–4] (free) populations were analyzed by a two-proportion z test. P-values are shown above the histogram bars. Arrows indicate bridge positions. The overviews in B and C and the bridges in E and F are presented in 3D in Videos 2 and 3, respectively. (D–F) Liposome surfaces are indicated by white dashed lines.
Cross-linking of CEACAM1 D(1–4)-His and D(2–4)-His. (A and C) Purified D(1–4)-His (A) or D(2–4)-His (C) were cross-linked with BS3 in the presence or absence of 2.5 mM Ca2+, 2.5 mM Mg2+, 3 mM EDTA, and Ni-NTA liposomes at a high protein/lipid ratio (hp; 1:10 [wt/wt]) or a low protein/lipid ratio (lp; 1:90 [wt/wt]) in various combinations. The samples were analyzed by Western blotting for CEACAM1 monomers/dimers/multimers. The black line indicates that intervening lanes have been spliced out. (B) Quantification of D(1–4)-His dimers and higher order multimers. (D) Quantification of D(2–4)-His dimers. Data show mean ± standard deviation of three independent experiments. Statistical analyses (t test) were made with reference to cross-linked protein with no extra additions (Untr.). Significant p-values (α level 0.05) are shown above the histogram bars.

Cross-linking of free and liposome-bound CEACAM1 ectodomains

To learn more about the homophilic binding interactions of the rat CEACAM1 ectodomains and possible adhesion-induced intracluster reorganization, we performed chemical cross-linking of D(1–4) and D(2–4) with BS3 (bis[sulfosuccinimidyl] suberate), which has a length of 11.4 Å between the amine-reactive groups. Cross-linking of D(1–4) in solution resulted in stabilization of dimers, trimers, tetramers, and higher multimers (Fig. 7, A and B). Addition of Ca/Mg ions or EGTA did not significantly change the proportion of the different cross-linked species, but EDTA decreased the abundance of cross-linked trimers/tetramers/multimers. Cross-linking of D(2–4) also resulted in stabilization of dimers, but the percentage of dimers was fourfold lower than for D(1–4) (Fig. 7, C and D). Oligomer species higher than dimers were not observed, and Ca/Mg or EDTA did not influence the level of dimerization.

Cross-linking of liposome-associated CEACAM1 ectodomains at high D(1–4)/liposome ratio (protein/lipid [wt/wt] 1:10) revealed a significant increase in the level of dimers and a decrease of higher order oligomers that was not statistically significant (P = 0.123; Fig. 7, A and B). At low D(1–4)/liposome ratio (protein/lipid [wt/wt] 1:90), the level of the D(1–4) dimers was similar to that of nonliposome-bound CEACAM1, whereas the higher oligomers disappeared almost completely (Fig. 7, A and B). The amount of cross-linkable D(2–4) dimers was not changed by liposome anchoring either in the presence or absence of Ca/Mg (Fig. 7, C and D). These...
results demonstrate that the more complex oligomerization of D(1–4) is a function of the interactions of the D1 domain and suggest that the adhesion-promoting property of CEACAM1 is an important regulator of the organization and intermolecular interactions of CEACAM1 within the membrane-bound clusters.

**Discussion**

In this study, we show that CEACAM1 behaves as a molecular system characterized by dynamic homophilic binding interactions. Three different methods, SPR-binding analysis, molecular tomography, and chemical cross-linking, consistently demonstrated that CEACAM1 ectodomains occur as a mixture of monomers, dimers, and oligomers. The tomographic analyses showed that the CEACAM1 ectodomain is flexible, being able to adopt several different conformations as a result of hinge regions between all of the Ig domains. Antiparallel trans-dimers (C-dimers) and parallel cis-dimers (A-dimers) could be distinguished. Also, the SPR-binding analyses identified the C- and A-dimerization reactions and demonstrated a rapid transition between monomers and C-dimers. The N-terminal D1 domain participated both in C- and A-dimerization, whereas domains D2–D4 were involved only in A-dimerization. All three methods revealed a partial dependence of divalent cations, which favored decreased ectodomain flexibility and enhanced formation of multimeric complexes. Because both dimerization reactions were enhanced, extracellular Ca/Mg should contribute to formation of CEACAM1-mediated cell adhesion.

Although homophilically binding proteins have been analyzed by flow cell biosensor techniques (Korotkova et al., 2008), it has previously not been possible to analyze the kinetics of such a system because no algorithms have existed that can deal with a situation in which the same binding interactions take place both within the mobile phase and between the mobile and solid phases. However, the novel curve-fitting algorithms that we present in this study allowed us to identify two different reactions that resulted in the formation of two types of dimers and one type of trimer in the presence of divalent cations. The equilibrium dissociation constants that were obtained from the curve fitting were in good agreement with the equilibrium dissociation constants determined from the tomography data. Both methods gave constants that were of the same order of magnitude for either type of dimerization reaction and showed the same dependence on divalent cations. These findings strongly support the interpretations of both the SPR-binding analyses and the molecular electron tomography.

A closer analysis of the data showed that the tomography gave a 1.7-fold higher degree of monomers and a 0.7-fold lower degree of total dimers than the SPR-binding experiments in the presence of Ca/Mg. In addition, the differences in the equilibrium constants and the proportion of C- and A-dimers obtained by the two methods suggest that the proportion of A-dimers might have been overestimated in the tomography experiments and/or that the formation of A-dimers might have been underestimated in the SPR-binding determinations. Overestimation of the proportion of A-dimers in the tomograms would mean that some of the parallel dimers that we classified as A-dimers were in fact C-dimers in which the physical binding only was mediated by domain D1. Underestimation of A-dimer formation in the SPR-binding experiments could be caused by homophilic interactions between the immobilized CEACAM1 ectodomains that we could not account for in our SPR-binding algorithms. Another contributing factor might be that the formation of A-dimers, which we treated as one single reaction, is in reality composed of four cooperative reactions in which all of the four Ig domains participate.

The most surprising result was the extent of flexibility of the CEACAM1 ectodomain, which is not in agreement with predictions from x-ray crystallography of murine CEACAM1 (Tan et al., 2002) or x-ray and neutron scattering of human CEACAM5 (Boehm et al., 1996). This demonstrates that molecular electron tomography has some advantages in characterization of gross conformations of single native molecules under physiological conditions. CEACAM1 anchored to liposomal membranes also appeared as flexible monomers and dimers, but the majority occurred in microclusters of closely packed molecules. Such clustering is expected, as predicted by Grasberger et al. (1986), who demonstrated that the likelihood of forming dimers and higher oligomers increases many orders of magnitude when dimerizing proteins are oriented on membrane surfaces. The degree of clustering did not change significantly when the CEACAM1 ectodomain to lipid ratio was varied over a 90-fold range, which demonstrates that the clustering effect caused by orientation on a surface was much stronger than what would be expected from just increased protein concentration in liquid solution. In many of the clusters, it was not possible to distinguish the individual molecules or Ig domains. This is what is expected at a 20-Å resolution of closely packed molecules, but it should be emphasized that the cluster structures really reflect what assembled molecules look like at this level of resolution. However, in spite of this limitation, it was possible to calculate the approximate number of molecules within each cluster from determination of the cluster volumes.

Tomography showed that CEACAM1-induced liposome adhesion was mediated by reciprocal binding between D1 domains presented both by single molecules and by small clusters of CEACAM1. The bridging clusters were both smaller and had different molecular packing compared with the free-surface clusters, and the cross-linking results indicated that this was a result of formation of D1-mediated antiparallel adhesion. An important piece of information for the interpretation of the cross-linking data is that we could show that antiparallel C-dimers mediating cell–cell adhesion are not stabilized by amine-reactive cross-linkers (see Müller et al. on p. 569 of this issue). Thus, the cross-linked dimers should exclusively represent parallel A-dimers. Also, the truncated D(2–4) ectodomains gave rise to a small proportion of cross-linked dimers, although the weaker interaction mediated by D(2–4) was not picked up by the SPR analysis. However, the A-dimerization contributed by Ig domain D1 can also be stabilized by cross-linking because the proportion of stabilized
D(1–4) dimers was significantly larger than that of D(2–4) under all conditions. Importantly, these findings show that the D1 domain can participate in mutual cis-binding in addition to trans-binding, demonstrating that it has two different homophilic binding sites, which is in agreement with the interpretation of the trimers identified by electron tomography. The cross-linked multimers formed by the D(1–4) ectodomains are clearly caused by the presence of domain D1 because no complexes larger than dimers were formed by D(2–4) under any conditions. These multimers most likely form because of the flexibility of the ectodomains, which allows combinations of parallel, A-type interactions between Ig domains belonging to several D(1–4) ectodomains (Fig. 8).

The multimers that were formed by D(1–4) in solution disappeared when the protein was attached to liposomes at a low protein/liposome ratio. This can be explained by a dilution of the ectodomains on the membrane surface, which will result in persistence only of the most stable dimer forms in which all four Ig domains are in register with each other. At a higher protein/liposome ratio, a significant increase in the amount of cross-linked D(1–4) A-dimers occurred. This was caused by the presence of domain D1 in the CEACAM1 ectodomain but was not a result of increased local concentration of membrane-attached protein because a corresponding increase of cross-linked D(2–4) A-dimers did not occur. Instead, the enhanced A-dimer formation was most likely a function of D1-mediated, antiparallel (trans) C-dimerization because it only occurred at the higher protein/liposome ratio, which resulted in more efficient homophilic liposome adhesion. It was not simply a result of increased local concentration of CEACAM1 ectodomains in the membrane contact regions because the statistical analysis showed that the CEACAM1 clusters were at the same time rearranged to become significantly smaller. Rather, it suggests that formation of trans–C-dimers between liposomal membranes induced cis–A-dimer formation by an allosteric mechanism transmitted by Ig domain D1, which is in agreement with the tomographic results that showed an increased abundance of dimeric clusters in the liposomal bridges.

The present data highlight the central role of the N-terminal D1 Ig domain in the functional activity of CEACAM1. A crucial finding was that two different sites in the D1 domain could participate simultaneously in homophilic binding. It is well established that either the CFG or the ABED face can participate in dimer formation in Ig domains (Edmundson et al., 1975). However, a simultaneous interaction of two areas in the same Ig domain has, to our knowledge, only previously been suggested for the adhesion molecules JAM-1 (Kostrewa et al., 2001), NCAM (Soroka et al., 2003), and TAG-1 (Mörtl et al., 2007). The binding site in the CEACAM1 D1 domain that mediates trans-homophilic binding, and presumably operates in C-dimer formation, has been localized to the CFG face by site-directed mutagenesis (Watt et al., 2001). The exact location of the other binding site in relation to the CFG face is unknown and could not be determined from the electron tomograms. Although crystallographic structures showing domains in close contact do not necessarily reflect a biologically relevant association, the finding of two unglycosylated CEACAM1 D1 domains in the asymmetric crystal unit that interact hydrophobically across the ABED faces (Fig. 3 C; Fedarovich et al., 2006) might indicate that this could represent the second binding site. However, in the crystal structure of mouse CEACAM1, this surface was covered by the oligosaccharide N linked to Asn-70 (Tan et al., 2002), and therefore, it has been argued that the ABED surface of CEACAM N-terminal domains cannot participate in physiological homophilic binding (Fedarovich et al., 2006). However, such a scenario should not be completely ruled out because oligosaccharides are flexible units, and it might be possible that the ABED surface would transiently and dynamically expose for participation in protein–protein interactions. Because CEACAM1 is variably glycosylated in different cell types and tissues (Odin et al., 1988; Singer et al., 2002), glycosylation might even represent a mechanism to control its putative potential to form A-dimers.

An important feature of N-terminal D1 domains in the CEACAM family is that they, in contrast to other Ig domains, lack intradomain disulfide bridges (Watt et al., 2001). This might be a prerequisite for the allosterically regulated homophilic binding interactions that we propose play an important role in CEACAM1 adhesion–induced transmembrane signaling. A similar lack of disulfides occurs in the N-terminal Ig domain of CD2, which also has been demonstrated to change conformation upon adhesion (Li et al., 1996). Interestingly, the N-terminal domain of CD2 has been shown to be in a metastable state, which can result in exchange of β strands between the domains in a CD2-dimer, resulting in a dimeric structure in which each domain is formed by the intercalation of two polypeptide
chains (Murray et al., 1995). If this would also occur in the CEACAM1 N-terminal domains, it might be the basis for the allosterism they transmit.

Allosterism has been demonstrated to occur in other CAMs such as selectins (Springer, 2009; Waldron and Springer, 2009) and integrins (Xiao et al., 2004). Integrin allosterism, which is important for bidirectional transmembrane signaling, is transmitted through the extracellular heterodimeric domains and is coupled to alteration of the intermolecular distance between the cytoplasmic domains of the α and β chains (Kim et al., 2003; Xiao et al., 2004). Because of the heterodimeric nature of integrins, the allostery can operate at the single molecule level. This is in contrast to the single chain molecule CEACAM1, for which the proposed allosteric changes instead cause alterations of the dimer/oligomer assembly.

Adhesion-induced, altered lateral organization of CEACAM1 in the adhesion bridges, manifested as increased parallel dimer and smaller clusters, has important implications for the transmembrane signaling by CEACAM1 and suggests a mechanism for how homophilic CEACAM1-mediated cell–cell adhesion can influence intracellular signaling. The altered organization of the ectodomains would be transmitted via the transmembrane domains, resulting in a corresponding alteration of the intermolecular organization of the cytoplasmic domains (Fig. 8). This might in turn influence binding/activation of SH2 domain–carrying enzymes, such as c-Src, SHP-1, and SHP-2, to the tyrosine-phosphorylated cytoplasmic domains of CEACAM1-L, causing a shift in the balance of kinase/phosphatase activation. In the accompanying paper (Müller et al., 2009), we demonstrate that such a mechanism involving adhesion-influenced dynamic changes of CEACAM1 microcluster organization indeed exists in epithelial cells.

Materials and methods

Cloning and expression of fc- and His-tagged proteins

For the rat and human D(1–4)-Fc constructs, rat and human CEACAM1 ectodomains (GenBank accession no. J04963 and X16354) and human IgG Fc (GenBank accession no. BC014258) were amplified (rat CEACAM1, 5′-AAGCTTGGATCCGCCGCGGGTGATATGCTG-3′ and 5′-GAATTCAGAATTCTTCTGTTGATGAGAAGCG-3′; human CEACAM1, 5′-AAGCTTGGATCCGCCGCGGGTGATATGCTG-3′; human CEACAM1, 5′-GAATTCAGAATTCTTCTGTTGATGAGAAGCG-3′; and human Fc, 5′-GAATTCATGGGACCTCTCAGCC-3′; and 5′-GAATTCAGAATTCTTCTGTTGATGAGAAGCG-3′). In the resulting vector, the amplified rat CEACAM1 ectodomain (annealed 5′-CATGGGGCACCTCTCAGCC-3′ and 5′-CATGGGGCACCTCTCAGCC-3′), the amplified human CEACAM1 ectodomain (annealed 5′-AGCTGACTATGATGATGACCGGTGCCCTGAAAATAAAGATTCTCGCAG-3′ and 5′-AGCTGACTATGATGATGACCGGTGCCCTGAAAATAAAGATTCTCGCAG-3′), and the amplified human Fc (annealed 5′-GATCCCGGGCAGCTGGAGAATCTTTATTTTCAGGGCA-3′) were ligated sequentially into the HindIII–EcoRI–XhoI sites of pcDNA3.1 (+) (Invitrogen). For the rat D(1–4)-His construct, the BsrGI-AgeI cassette of pcDest 40 V5-His (Invitrogen) was replaced with a duplex encompassing the TEV protease cleavage site (annealed 5′-GATCCCGGGCAGCTGGAGAATCTTTATTTTCAGGGCA-3′ and 5′-GATCCCGGGCAGCTGGAGAATCTTTATTTTCAGGGCA-3′) and ligated sequentially into the HindIII–EcoRI–XhoI sites of pcDNA3.1 (+) (Invitrogen). For the rat D(1–4)-His construct, the BsrGI-AgeI cassette of pcDest 40 V5-His (Invitrogen) was replaced with a duplex encoding the TEV protease cleavage site (annealed 5′-GTACAAAGCTTAAGCTG-3′ and 5′-GTACAAAGCTTAAGCTG-3′). In all three models, the bulk analyte is referred to as Abulk (given in grams/liter). The different forms of the analyte, i.e., monomer, dimer, and trimer, are referred to as [A], [A]2, and [A]3, respectively (given in micromolars/liter). To ease the reading of the coupled differential equations applied for numerical integration (Eqs. 1.3–1.5, 2.3–2.8, and 3.6–3.15), the brackets have been dropped, and [A], [A]2, and [A]3 are used as A, A2, and A3, respectively. Bulk concentrations of these forms are indicated with the subscript B. Disassociation rate constants (kD) are given in units of liters/(micromolars × seconds). Association rate constants (kA) are given in liters/micromolars. The bulk analyte concentration is kept constant by a continuous injection at a constant flow rate. Thus, in our models, the bulk analyte forms (i.e., monomer, dimers, and trimers) are assumed to be in equilibrium. During sample injection, the amount of unoccupied ligand, i.e., [L], decreases with time until equilibrium is reached. Thus, the concentration of free-ligand sites, [L], at a particular moment, is given by [L] = [L]0 − [A]. where [L]0 is the total amount of accessible surface-immobilized ligand [i.e., available binding sites] at time 0, and [A] is the amount of complex formed at time t of analyte injection. The response signal, R, measured by the SPR-based sensor, is proportional to the amount of complex formed at the detector surface multiplied by the factor MwG, i.e., R = MwG[A], where Mw is the molecular mass of the analyte, and G is a factor converting the concentration to R values (G = 10,000 R cm²/g of protein; Steenberg et al., 1991). This allows for the rate equations to be expressed in R units. Rmax is defined as the theoretical maximum-binding response at surface saturation. All three models...
assume a low ligand density and do not account for interactions or bridge formations between ligand units.

The models were programmed in IGOR Pro (version 6.03A2) and were applied together with the IGOR Pro Global Fit Procedure based on a nonlinear least-squares method using the Levenberg-Marquard algorithm (Press et al., 1999). For numerical integration, we applied the IGOR Pro algorithm based on a fifth order Runge-Kutta-Fehlberg method (Press et al., 1999).

1. Formation of one class of dimers by one homophilic binding site. This model describes the simplest system of homophilic binding interactions,

\[ A + A \leftrightarrow A_2 \]

where \( A \) and \( A_2 \) define the monomer and dimer forms of the analyte, respectively. At all conditions, it holds for the analyte in the units of grams/liter that

\[ A_2 + A + A_{\text{Total}} = 0. \tag{1.1} \]

At equilibrium, it holds that \([A]^2 k_2 = [A_2] k_o\), where \([A]\) and \([A_2]\) represent the monomer concentrations. Thus, \([A_2] = [A]^2 / k_o\), where the equilibrium dissociation constant \( k_o = k_o/k_a \). From Eq. 1.1, it follows that

\[ \frac{2M_w}{K_D} [A]^2 + M_w [A] - A_{\text{Total}} = 0. \]

For simplicity, this quadratic equation can be rearranged to the classical form \( a x^2 + b x + c = 0 \), defined as \( a = 2M_w / k_o \), \( b = M_w \), \( c = -1 \times A_{\text{Total}} \), and \( x = [A] \), with the relevant solution

\[ [A] = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{1.2} \]

In addition to Eq. 1.2 and the relation \([A_2] = [A]^2 / k_o\), which provide the bulk concentrations of the monomer and dimer forms for a given \( A_{\text{Total}} \), the curve-fit model is composed of the following coupled differential equations, applied for numerical integration:

\[ \frac{dA}{dt} = \frac{k_1}{h} (A^2 - A) - k_p A \left( A + \frac{R_{\text{max}} - R}{hM_wG} \right) + k_d \left( 2A + \frac{R}{hM_wG} \right) \tag{1.3} \]

\[ \frac{dA_2}{dt} = \frac{k_2}{h} (A^2 - A_2) + k_p A^2 - k_d A_2, \tag{1.4} \]

\[ \frac{dR}{dt} = k_p A (R_{\text{max}} - R) - k_d R. \tag{1.5} \]

2. Formation of two classes of dimers (types 1 and 2) by two different homophilic binding sites. By the introduction of a second binding site in the monomer, two more variables, \( k_1 \) and \( k_2 \), need to be accounted for. At all conditions, it holds for the analyte (given in grams/liter) that

\[ A_{\text{Total}} = 0. \tag{2.1} \]

which at equilibrium gives

\[ \left( \frac{2M_w}{K_{D1}} \right) [A]^2 + M_w [A] - A_{\text{Total}} = 0, \]

with the following solution for the molar concentration of the monomer form:

\[ [A] = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \tag{2.2} \]

In addition to Eq. 2.2 and the relations \([A_{12}] = [A]^2 / k_o\), and \([A_{22}] = [A]^2 / k_{o2}\), which provide the bulk concentrations of the monomer and dimer forms for a given \( A_{\text{Total}} \), the curve-fit model is composed of the following six coupled differential equations, applied for numerical integration (surface-bound type 1 and type 2 dimers are indicated with \( R_{\text{1d}} \) and \( R_{\text{2d}} \), respectively):

\[ \frac{dA_1}{dt} = \frac{k_1}{h} (A^2 - A_{12}) + k_d A^2 - k_d A_{12}, \tag{2.3} \]

\[ \frac{dA_2}{dt} = \frac{k_2}{h} (A^2 - A_{22}) + k_d A^2 - k_d A_{22}, \tag{2.4} \]

\[ \frac{dA_{1d}}{dt} = \frac{k_1}{h} (A_{1d}^2 - A_{12}) + k_d A^2 - k_d A_{12}, \tag{2.5} \]

\[ \frac{dA_{2d}}{dt} = \frac{k_2}{h} (A_{2d}^2 - A_{22}) + k_d A^2 - k_d A_{22}, \tag{2.6} \]

\[ \frac{dR_{1d}}{dt} = k_p A (R_{\text{max}} - R_{1d}) - k_d R_{1d}, \tag{2.7} \]

\[ \frac{dR_{2d}}{dt} = k_p A (R_{\text{max}} - R_{2d}) - k_d R_{2d}, \tag{2.8} \]

3. Formation of two classes of dimers (types 1 and 2) and one class of trimers by two different homophilic binding sites. This model is a modification of the model described in section 2. If both binding sites participate, larger complexes can be formed. If we set the limit for a larger complex assembly at the trimer level, Eq. 2.1 expands to

\[ A_1 + A_{12} + A_2 + A - A_{\text{Total}} = 0. \tag{3.1} \]

In this case, addition of the trimmer term does not require additional fit parameters. At equilibrium, it holds that

\[ [A_{12}] = \frac{k_{12}}{k_1} [A]^2, \tag{3.2} \]

\[ [A_{22}] = \frac{k_{22}}{k_2} [A]^2, \tag{3.3} \]

\[ [A_3] = \frac{k_1 k_2}{k_{12} k_{22}} [A]^3. \tag{3.4} \]

Recalling Eq. 3.1, the total amount of analyte is composed of monomer, dimer, and trimer forms (i.e., \( A_{\text{Total}} \) in grams/liter). It is beneficial to express these forms with simple functions of the monomer \([A]\) combined with rate constants to provide the minimal amount of variables. We define \( M_w \) as the molecular mass of the monomer and place it with the proper multiplication terms, i.e., times three for a trimer and times two for the dimers, together with Eqs. 3.2–3.4 into Eq. 3.1 to give

\[ M_w \left( \frac{3 k_{12} k_{22}}{k_1 k_{22}} [A]^3 + \frac{2 k_1 k_{12} + k_2 k_{22}}{k_1 k_2} [A]^2 + [A] \right) - A_{\text{Total}} = 0. \]

This equation can be rearranged to the general form of the cubic equation:

\[ \alpha x^3 + \beta x^2 + \gamma x + \delta = 0, \]

where

\[ \alpha = \frac{3M_w k_{12} k_{22}}{k_1 k_{22}}, \]

\[ \beta = \frac{2M_w k_{12} + \frac{2M_w k_{12} + \frac{k_2}{k_1}}{k_1 k_2} [A]^2 + [A]}, \]

\[ \gamma = \frac{-M_w k_{12} k_{22} + \frac{2M_w k_{12} + \frac{k_2}{k_1}}{k_1 k_2}}{k_1 k_2}, \]

\[ \delta = \frac{-M_w k_{12} k_{22} + \frac{2M_w k_{12} + \frac{k_2}{k_1}}{k_1 k_2} [A]}{k_1 k_2}. \]
\[ \beta = \frac{2M_W(k_{\alpha\delta}A_2 + k_{\alpha\gamma}A_1)}{k_{\alpha\delta}A_2}. \]

\[ \gamma = M_W. \]

\[ \delta = -A_{\text{total}} \text{ and} \]

\[ x = [A]. \]

Of the three solutions to the cubic equation, two involve complex numbers, and therefore, with regard to our model, have no physical relevance. The third one is a real number solution, which is relevant for addressing the monomer concentration in the sample (i.e., analyte), yielding (3.4)

\[ \frac{d\beta}{dt} = \frac{1}{hM_WG} \left(2M_WGk_{\alpha\delta} \left(A_{\beta\gamma} - A_{\alpha\gamma}\right) + hM_WG \left(k_{\alpha\delta}A^2 - (k_{\alpha\gamma} + k_{\alpha\gamma})A_{\beta\gamma} + k_{\alpha\gamma}A_1\right) \right), \]

\[ +k_{\alpha\gamma}R_{\alpha\beta\gamma} + k_{\alpha\gamma}A_{\beta\gamma} \left(-R_{\max} + 2(R_{\alpha\beta} + R_{\alpha\gamma}) + R_{\alpha\gamma} + R_{\alpha\beta\gamma} + R_{\alpha\gamma}\right). \]

\[ \frac{dA_3}{dt} = \frac{k_{\alpha\gamma}}{h} \left(A_{\alpha\gamma} - A_3\right) + A(k_{\alpha\gamma}A_{\beta\gamma} + k_{\alpha\gamma}A_{\alpha\gamma}) - A_3(k_{\alpha\gamma} + k_{\alpha\gamma}). \]

(3.10)

\[ \frac{dR_{\alpha\gamma}}{dt} = \frac{1}{hM_WG}(2M_WGk_{\alpha\delta} \left(A_{\beta\gamma} - A_{\alpha\gamma}\right) + hM_WG \left(k_{\alpha\delta}A^2 - (k_{\alpha\gamma} + k_{\alpha\gamma})A_{\beta\gamma} + k_{\alpha\gamma}A_1\right) \right), \]

\[ +k_{\alpha\gamma}R_{\alpha\beta\gamma} + k_{\alpha\gamma}A_{\beta\gamma} \left(-R_{\max} + 2(R_{\alpha\beta} + R_{\alpha\gamma}) + R_{\alpha\gamma} + R_{\alpha\beta\gamma} + R_{\alpha\gamma}\right). \]

However, this solution is rather fragile if applied on a real number format to curve-fit algorithms because of somewhat complicated variable compositions within the root terms. For this reason, subfunctions for the trimer model based on the fact that, in contrast to a single form of trimer in solution, the surface-bound trimers cannot be treated as equal. Because one of the three components belonging to the trimer will be the immobilized ligand, an asymmetry is introduced. Because of different responses for the association or dissociation of monomers versus dimers to a surface-bound ligand, together with different interaction rates for the two types of binding sites involved, the model has to account for two types of dimers and three types of trimers contributing to the SPR-based signal. These ligand complexes are distinguished as \( R_{\alpha\beta}, R_{\alpha\gamma}, R_{\alpha\beta\gamma}, \) and \( R_{\alpha\gamma} \) (Fig. S1). As for the other two models, this model has to be provided with information on \( A_{\text{total}} \) (in grams/liter), \( M_W, R_{\alpha\beta}, \) and initial values for the rate constants. Finally, together with Eq. 3.5 for the estimation of \([A^0]\), Eqs. 3.2 and 3.3 for the estimation of the dimer \([A^0 \gamma]\) form and Eq. 3.4 for the estimation of the \([A^0 \beta\gamma]\) form, the curve-fit model is composed of the following 10 coupled differential equations, applied for numerical integration:

\[ \frac{dA_3}{dt} = \frac{k_{\alpha\gamma}}{h} \left(A_{\alpha\gamma} - A_3\right) + A(k_{\alpha\gamma}A_{\beta\gamma} + k_{\alpha\gamma}A_{\alpha\gamma}) - A_3(k_{\alpha\gamma} + k_{\alpha\gamma}). \]

(3.9)

\[ \frac{dR_{\alpha\gamma}}{dt} = \frac{1}{hM_WG}(2M_WGk_{\alpha\delta} \left(A_{\beta\gamma} - A_{\alpha\gamma}\right) + hM_WG \left(k_{\alpha\delta}A^2 - (k_{\alpha\gamma} + k_{\alpha\gamma})A_{\beta\gamma} + k_{\alpha\gamma}A_1\right) \right), \]

\[ +k_{\alpha\gamma}R_{\alpha\beta\gamma} + k_{\alpha\gamma}A_{\beta\gamma} \left(-R_{\max} + 2(R_{\alpha\beta} + R_{\alpha\gamma}) + R_{\alpha\gamma} + R_{\alpha\beta\gamma} + R_{\alpha\gamma}\right). \]

(3.11)

\[ \frac{dR_{\alpha\beta\gamma}}{dt} = k_{\alpha\gamma}R_{\alpha\gamma}A - (k_{\alpha\gamma} + k_{\alpha\gamma})R_{\alpha\beta\gamma}. \]

(3.12)

\[ \frac{dR_{\alpha\beta\gamma}}{dt} = k_{\alpha\gamma}R_{\alpha\beta\gamma} - \left(2R_{\alpha\beta} + R_{\alpha\gamma}\right) + R_{\alpha\gamma} + R_{\alpha\beta\gamma} + R_{\alpha\gamma}. \]

(3.13)

\[ \frac{dR_{\alpha\beta}}{dt} = A(k_{\alpha\gamma}R_{\alpha\beta} + k_{\alpha\gamma}R_{\alpha\gamma}) - R_{\alpha\beta}(k_{\alpha\gamma} + k_{\alpha\gamma}). \]

(3.14)

\[ \frac{dR_{\alpha\beta\gamma}}{dt} = \frac{dR_{\alpha\gamma}}{dt} + \frac{dR_{\alpha\beta}}{dt} - \frac{dR_{\alpha\beta\gamma}}{dt} + \frac{dR_{\alpha\gamma}}{dt} \]

(3.15)

**Molecular electron tomography**

1.0–2.9 mg/mL rat D(1–4)-His and rat D(2–4)-His in HBS (150 mM NaCl and 10 mM Hepes, pH 7.4) supplemented with either 2 mM CaCl₂, 2 mM MgCl₂, or 3 mM EDTA were mixed with a suspension of washed colloidal 10-nm BSA-coated gold markers (AuroProbe EM protein A G10; GE Healthcare), placed on carbon-coated grids (Quantifoil R1.2/1.3; Quantifoil), and plunged into liquid ethane. Liposome-anchored, His-tagged proteins were prepared by mixing the protein solution with a suspension of washed colloidal 10-nm liposomes before the addition of colloidal gold. A field emission–scanning transmission electron microscope (200-kV acceleration voltage; CEM 200; Philips) equipped with a detection system with a magnification of 41,800 on a 2,048 x 2,048-pixel charge-coupled device chip with a raster size of 24 µm, giving a pixel size of 5.74 Å at 1° intervals. The total dose on the cryosamples device chip with a raster size of 24 µm, giving a pixel size of 5.74 Å/

\[ \frac{dt}{\text{day}} = \frac{1}{hM_WG}(2M_WGk_{\alpha\delta} \left(A_{\beta\gamma} - A_{\alpha\gamma}\right) + hM_WG \left(k_{\alpha\delta}A^2 - (k_{\alpha\gamma} + k_{\alpha\gamma})A_{\beta\gamma} + k_{\alpha\gamma}A_1\right) \right), \]

\[ +k_{\alpha\gamma}R_{\alpha\beta\gamma} + k_{\alpha\gamma}A_{\beta\gamma} \left(-R_{\max} + 2(R_{\alpha\beta} + R_{\alpha\gamma}) + R_{\alpha\gamma} + R_{\alpha\beta\gamma} + R_{\alpha\gamma}\right). \]

(3.10)
were initially 3D reconstructed using the filtered backprojection principle and were subsequently refined by COMET (version 4.5.7; Skoglund et al., 1996; Sandin et al., 2004; Rullgård et al., 2007) to improve the signal to noise ratio. Physical parameters describing the sample and used during data collection were used in the COMET regularization. These entailed a 1-μm underfocus, an estimated amplitude contrast value of −0.10 to 0.15 as defined previously (Fanelli and Öktem, 2008), a lower density histogram truncation parameter of 0.03–0.05, and a regularization parameter of 0.6 in the span of 0.0–1.0 in which 1.0 means no entropy regularization. The prior density resolution was kept at 80–100Å resolution. Two different segmentation methods were used for identifying molecular structures: (1) gray-level thresholding applying BOB software (provided by K. Chin-Purcell, Minnesota Supercomputing Institute, Minneapolis, MN) and (2) SWS. In segmentation based on gray-level thresholding, all voxels having gray levels above a certain threshold are considered to belong to the particle, and voxels with lower gray values are regarded as background. In watershed segmentation (Beucher, S., and C. Lantuejoul. 1979. International Workshop on Image Processing), the gray-level image is considered as a topographic map, and the final segmentation corresponds to the catchment basins, with one for each local minimum. When the propagation fronts from different minima meet, a watershed is built to prevent further region growing. Because many minima correspond to small, nonrelevant gray-level variations, which results in an oversegmentation, a slightly modified version of watershed segmentation was used, namely SWS (Vincent, 1993).

In this case, region growing is only allowed from a set of seeds, which can be, e.g., a limited set of local minima. We used SWS applied to the gradient magnitude information extracted from the 3D reconstruction. The algorithm to extract seeds, which was described previously by Vincent (1993), was adjusted to suit our application.

Image processing of Protein Data Bank (PDB) files

The crystal structures of human CEACAM5 ectodomain D1 (PDB accession no. 2GSG; Korokova et al., 2008), human CEACAM1 ectodomain D1 (PDB accession no. 2GK2; Fedarovich et al., 2006), and mouse CEACAM1 ectodomain (D1–D4 form; PDB accession no. 1H6Z; Tan et al., 2002) were downloaded from the PDB database. The RHOGENE program (available for academic use through U. Skoglund) was used to generate 3D map files of their theoretical nuclear electron scattering. The 3D maps were low-pass filtered to 20Å resolution, corresponding to the resolution of the electron tomograms, and visualized using BOB software. The PDB files were also used to create ribbon models and space-filling models with atomic resolution using the Molecular Biology Toolkit protein workshop or iMol (version 0.30; http://www.pirx.com/iMol/).

Liposome preparation

Unilamellar, mixed-lipid liposomes with a 90:5:5 molar composition of phosphatidylcholine/phosphatidic acid/DOGS-NTA-Ni were prepared with lipids obtained from Avanti Polar Lipids, Inc. Chloroform solutions of the lipids (10 mg total lipid) were mixed, evaporated under a stream of nitrogen, and dissolved in 1 ml OG/HBS (20 mg/ml octyl glucopyranoside in 150 mM NaCl and 10 mM Hepes, pH 7.4). The lipid/OG solution was dialyzed (Spectra/Por dialysis tubing, molecular weight cut off of 12–14,000) at room temperature against 5 × 600 ml HBS and stored at 4°C. The lipid/OG solution was subsequently refined by COMET (version 4.5.7; Skoglund et al., 1996; Rullgård et al., 2007) to improve the signal to noise ratio. Images were imported into Photoshop (Adobe) for processing.

Turbidity measurements

Protopotoliposomes were prepared by adding different concentrations of rat CEACAM1 D(1–4)-His or D(2–4)-His ectodomains to a constant amount of unlabeled Ni-NTA liposomes (final concentration of 1 mg protein/ml in 1 ml OG/HBS). The liposomes were mixed, incubated for 30 min with the chemical cross-linker BS3 (Thermo Fisher Scientific) at a final concentration of 2.5 mM, and the samples were subsequently incubated with 20 mM Tris, pH 8, for 20 min. Untreated, non-cross-linked proteins were not incubated for 1 h at room temperature with various concentrations of Ni-NTA liposomes to yield protopotoliposomes with protein/lipid ratios of 1:10 to 1:90 [wt/wt]. The turbidity was recorded as optical density at 595 nm at various times after protein/ligand mixing.

Cross-linking experiments

Samples of rat CEACAM1 D(1–4)-His or D(2–4)-His ectodomains (final concentrations of 0.1 mg protein/ml) in plain HBS were supplemented with 2.5 mM Ca2+, 2.5 mM Mg2+, 3 mM EDTA, and 0.3 mM Ca2+, 2.5 mM Mg2+, 3 mM EGTA or were not incubated for 1 h at room temperature with various concentrations of Ni-NTA liposomes to yield protopotoliposomes with protein/lipid ratios of 1:10 and 1:90 [wt/wt], respectively. The different samples were incubated for 30 min with the chemical cross-linker BS3 (Thermo Fisher Scientific) at a final concentration of 2.5 mM, and the reactions were quenched with 20 mM Tris, pH 8, for 20 min. Untreated, non-cross-linked proteins were incubated with 20 mM Tris, pH 8, for 20 min served as controls. The samples were analyzed by SDS-PAGE under reducing conditions, and Western blotting was performed with rat CEACAM1-specific Mab Be9.2 (Becker et al., 1986) or Pab ccC16 (Singer et al., 2000) to detect D(1–4)-His and D(2–4)-His, respectively. Protein detection on Western blots using nitrocellulose membrane (Schleicher & Schüll) was performed using SuperSignal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific). Chemiluminescence was detected using a digital system (LAS-1000; Fujifilm). Quaternary structure was determined using Image Gauge software (Fujifilm). Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.200904149/DJC1.

References


Figure S1. Reaction schemes for homophilic binding interactions in a BIAcore flow cell. The same kind of molecules occur both as analytes in the soluble phase and as immobilized ligands in the solid phase. (A) Formation of one class of dimers by one homophilic binding site. (B) Formation of two classes of
Figure S2. SDS-PAGE of purified recombinant CEACAM1 ectodomains. Purified rat CEACAM1 D(1–4)-His, D(2–4)-His, D(1–4)-Fc, D(2–4)-Fc, and human CEACAM1 D(1–4)-Fc were separated on a NuPAGE Novex Bis-Tris gel (4–12%; Invitrogen) in NuPAGE-MES-SDS or NuPAGE-MOPS-SDS running buffer (Invitrogen) under reducing conditions and stained with Coomassie brilliant blue. Black lines indicate that intervening lanes have been spliced out.

dimers (types 1 and 2) by two different homophilic binding sites. (C) Formation of two classes of dimers (types 1 and 2) and one class of trimers by two different homophilic binding sites. The association and dissociation reactions underlying the different complex formations are characterized by a set of kinetic rate constants, $k_{a1}/k_{d1}$ and $k_{a2}/k_{d2}$, respectively. These binding reactions occur both between the soluble molecules in the fluid phase and with the immobilized molecules at the surface. In order for the surface interactions to take place, the soluble molecules in the bulk phase must be transported to the surface. This mass transport process is characterized by a mass transport coefficient, $k_c$. Because the monomers, dimers, and trimers are of different sizes, they have different mass transport coefficients, $k_{c1}$, $k_{c2}$, and $k_{c3}$, respectively. (C) In the trimer model, an asymmetry is introduced because the immobilized ligand makes up one of the three components of the trimer that is formed on the surface. Therefore, two types of dimers and three types of trimers contribute to the SPR-based signal. These complexes are the following: AL, type 1 dimer; LA, type 2 dimer; AAL, A-type 2-A-type 1-L; ALA, A-type 1-L-type 2-A; and LAA, L-type 2-A-type 1-A. A, analyte; L, ligand.
Video 1.  **D(1–4)-His molecules analyzed by molecular tomography.** 3D reconstructions of D(1–4)-His molecules presented and described in detail in Fig. 4 [A–H] are shown in the same order, rotating around one axis. Monomeric D(1–4)-His adopting an extended (Fig. 4, A and B), kinked (Fig. 4 C), or back-folded (Fig. 4 D) form. (Fig. 4 E) D(1–4)-His dimer, interacting exclusively via the D1 domains (C-dimer; see Results). (Fig. 4, F and G) D(1–4)-His dimers interacting via three (Fig. 4 F) or four (Fig. 4 G) of their Ig domains (A-dimers; see Results). (Fig. 4 H) D(1–4)-His trimer consisting of a monomer binding via its D1 domain to an A-dimer. The units of the coordinates are pixels, and the pixel size is 5.74 Å.

Video 2.  **Proteoliposomes analyzed by molecular tomography.** Free surfaces of Ni-NTA liposomes with attached D(1–4)-His or D(2–4)-His ectodomains presented in Fig. 6 (B and C) are shown in the same order, rotating around one axis. The units of the coordinates are pixels, and the pixel size is 5.74 Å.

Video 3.  **D(1–4)-His molecules bridging two liposomes analyzed by molecular tomography.** An overview and details of the monomeric bridge (C-dimer) presented in Fig. 6 E and the heptameric bridge presented in Fig. 6 F are shown. The units of the coordinates are pixels, and the pixel size is 5.74 Å.