**Structural Model and trans-Interaction of the Entire Ectodomain of the Olfactory Cell Adhesion Molecule**

Nikolaj Kulahin,1,2,3,4 Ole Kristensen,2 Kim K. Rasmussen,1,2 Lars Olsen,2 Patrik Rydberg,2 Bente Vestergaard,2 Jette S. Kastrup,2 Vladimir Berezin,1 Elisabeth Bock,1 Peter S. Walmad,1,* and Michael Gajhede2, *

1Protein Laboratory, Department of Neuroscience and Pharmacology, Faculty of Health Sciences, University of Copenhagen, DK-2200 Copenhagen N, Denmark
2Biostructural Research, Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, DK-2100 Copenhagen O, Denmark
3ENKAM Pharmaceuticals A/S, DK-2100 Copenhagen O, Denmark
4Present address: Receptor Systems Biology Laboratory, Hagedorn Research Institute, DK-2820 Gentofte, Denmark

*Correspondence: psw@sund.ku.dk (P.S.W.), mig@farma.ku.dk (M.G.)

DOI 10.1016/j.str.2010.12.014

**SUMMARY**

The ectodomain of olfactory cell adhesion molecule (OCAM/NCAM2/RNCAM) consists of five immunoglobulin (Ig) domains (IgI–V), followed by two fibronectin-type 3 (Fn3) domains (Fn3I–II). A complete structural model of the entire ectodomain of human OCAM has been assembled from crystal structures of six recombinant proteins corresponding to different regions of the ectodomain. The model is the longest experimentally based composite structural model of an entire IgCAM ectodomain. It displays an essentially linear arrangement of IgI–V, followed by bends between IgV and Fn3I and between Fn3I and Fn3II. Proteins containing IgI–IgII domains formed stable homodimers in solution and in crystals. Dimerization could be disrupted in vitro by mutations in the dimer interface region. In conjunction with the bent ectodomain conformation, which can position IgI–V parallel with the cell surface, the IgI–IgII dimerization enables OCAM-mediated trans-interactions with an intercellular distance of about 20 nm, which is consistent with that observed in synapses.

**INTRODUCTION**

Neural cell adhesion molecules (CAMs) have adhesive properties that are of major importance for the self-assembly of interconnections between cells in the mammalian brain. Additionally, they have many other functions related to cell-cell signaling (reviewed in Shapiro et al., 2007). CAMs are divided into several families, including the integrins, selectins, cadherins, and CAMs of the immunoglobulin (Ig) superfamily (Juliano, 2002). CAMs are generally multidomain proteins with the extracellular part comprising domains in a modular arrangement. The molecular mechanisms controlling the various functions of CAMs have attracted much attention from the field of drug discovery. For example, CAMs can be essential therapeutic targets for the treatment of cancer or neurodegenerative diseases (Raveh et al., 2009; Rice et al., 2005).

Mammalian forms of olfactory cell adhesion molecule (OCAM/NCAM2/RNCAM) were cloned in 1997 (Alenius and Bohm, 1997; Paoloni-Giacobino et al., 1997; Yoshihara et al., 1997). OCAM exists in a transmembrane and a glycosylphosphatidylinositol-anchored isoform resulting from alternative splicing of the transcript from a single gene (reviewed in Kulahin and Walmad, 2010). OCAM transcripts are found in several tissues (Paoloni-Giacobino et al., 1997) including different regions of the brain where it is believed to regulate neurite outgrowth, axonal guidance, synapse formation, and the formation of dendritic bundles (Hamlin et al., 2004; Ichinohe et al., 2003; Paoloni-Giacobino et al., 1997). The expression of OCAM has been suggested to potentially influence certain types of neurological diseases, including autism (Molloy et al., 2005) and Down’s syndrome (Paoloni-Giacobino et al., 1997), and studies also indicate that the expression of OCAM can be changed in, for example, prostate cancer cells (Edwards et al., 2005; Xu et al., 2001).

OCAM is a paralog of neural cell adhesion molecule 1 (NCAM1) (see Owczarek et al., 2009), and the extracellular part of both proteins consists of five N-terminal Ig domains (IgI–V), followed by two fibronectin type III domains (Fn3I–II) (Figure 1). Cell aggregation studies have shown that OCAM mediates homophilic trans-interactions, but the domains involved in the interaction have not been identified. Additionally, in vitro binding assays suggest that OCAM does not participate in heterophilic interactions with NCAM1 (Yoshihara et al., 1997).

The molecular basis for homophilic interactions within the IgCAMs has been investigated in, for example, the L1CAM family (Haspel and Grumet, 2003) and DSCAM (Sawaya et al., 2010). Within the NCAM family, the crystal structures of the N-terminal Ig domains IgI–II (Kasper et al., 2000) and IgI–III (Soroka et al., 2003) of NCAM1 have been determined (reviewed in Soroka et al., 2010), and recently, the crystal structure of OCAM Ig1 was published (Rasmussen et al., 2008). Structural models of the entire ectodomain only exist for relatively short IgCAMs, such as P0 (one Ig domain), JAM1 (two Ig...
domains), and SIRPα (three Ig domains) (Hatherley et al., 2009; Prota et al., 2003; Shapiro et al., 1996). For longer IgCAMs, structures of ectodomain segments are available for TAG-1 (four out of ten extracellular domains) and DSCAM (eight out of 16 extracellular domains) (Mortl et al., 2007; Sawaya et al., 2008). However, no structural models of the entire ectodomain of long IgCAMs have been proposed.

Here, we present a structural model of the entire ~700 amino acid-long ectodomain of human OCAM, assembled from crystal structures of six recombinant proteins with overlapping primary sequences. This represents the longest experimentally based composite structural model of the entire extracellular part of an IgCAM. The ectodomain forms homodimers via the membrane-distal IgI and IgII domains, resulting in a model for OCAM-mediated cell-cell adhesion with an intercellular distance of ~20 nm.

**RESULTS**

**Crystal Structures of the Extracellular Domains of OCAM**

The crystal structures of recombinant proteins corresponding to the IgI–II, IgI–III, IgI–IV, IgIV–F3I, and IgIV–F3II domains of human OCAM have been determined (Figures 2–4 and Table 1). The structures of IgI, IgII, and IgIII all fall into the intermediate 1 (I1) subset of Ig domains (Wang and Springer, 1998). Overall, these three domains are structurally similar to those of NCAM1, with root-mean-square deviations (rmsds) on Cα atoms ranging from 0.6 to 0.9 Å. However, important differences exist between the structures of the Ig domains from NCAM1 and OCAM. A C′ strand that is present in OCAM IgII (Figures 1 and 3) is missing in the corresponding NCAM1 IgII region, a region known to be involved in heparin binding (Kulahin et al., 2005b; Nielsen et al., 2010; Reyes et al., 1990). Consequently, the heparin-binding site of NCAM1 is not conserved between the two proteins (Figure 1). Moreover, the conformations of the loop regions corresponding to the NCAM1 heparin-binding sequence are very different between NCAM1 and OCAM, and
because the residues Glu153 and Glu154 (Uniprot code O15394) are disordered in the OCAM IgI–II, IgII–III, and IgI–III structures, this region may be more flexible in OCAM than in NCAM1.

The loop connecting the IgIII C and D strands is three residues shorter in OCAM than in NCAM1. Whereas this loop is disordered in the NCAM1 IgI–III structure, it is clearly defined in the OCAM IgI–III, IgII–III, and IgIII–IV structures. This ordering may be related to the hydrogen bond formed between Tyr257 N and Asn254 O. Notably, OCAM IgIV appears to belong to the variable (V) subset of Ig domains due to the presence of a putative C00 strand. However, whereas the C00 strand is well defined in the IgIII–IV structure, it is disordered in the IgIV–Fn3I and IgIV– Fn3II structures. The loop region 346-GDKS-349 is one of the most flexible regions in IgIV, and it is disordered in the OCAM IgIV–Fn3I and IgIV– Fn3II structures. This loop corresponds to the sequence in NCAM1 IgIV encoded by the VASE mini-exon (354-ASWTPRPEKQE-363) (Figure 1), which in NCAM1 is associated with downregulated neural plasticity (reviewed in Owczarek et al., 2009). The IgV structure is most closely related to the NMR structure of the same domain (PDB code 2kbg), with an rmsd of 1.1 Å.

Glycosylation of OCAM

Human OCAM contains eight potential sites for N-linked glycosylation, five of which are also found in NCAM1 (Figures 1 and 5). However, in contrast to NCAM1, OCAM has been reported not to be glycosylated with the unusual carbohydrate polysialic acid (PSA; long, negatively charged homopolymer of α2-8-N acetylated neuraminic acid). N-linked glycosylation was clearly visible at Asn177 in IgII, Asn219 in IgIII, Asn309 in IgIV, and Asn419, Asn445, and Asn474 in IgV and could be unambiguously fitted into the electron densities. The Asn177 glycosylation site is absent in NCAM1, due to the presence of lysine at this position. Glycosylation sites Asn445 and Asn474 in OCAM IgV, corresponding to the polysialylation sites in NCAM1, are shown in Figures 1, 4, and 5. Notably, the α helix between strands D and E in the NCAM1 Fn3I domain (562-AKEASMEG-569; Figure 1),
residues Arg224 and Gln301, this long strand may stabilize the protein formed in NCAM1 between Lys130 and Ala294. The OCAM IgIII bond between OCAM residues Lys127 and Ala287 is also present in NCAM1 (Glu128 and Lys293). The equivalent of the hydrogen bond between OCAM residues Glu125 and Lys127 is also seen in OCAM IgII–III. A salt bridge to the carbonyl oxygen atom of Ala287. The tilt angle in the NCAM1 structure, where the corresponding arginine forms an in-plane hydrogen bond through two water molecules. The observed tilt angle is 27°–28°, again corresponding to a near-linear arrangement of the domains.

The interdomain link between IgV and Fn3I in the OCAM IgV–Fn3I structure is more bent than the linker regions between Ig domains, with a tilt angle of 52°. One interdomain hydrogen bond between Trp413 and Gly524 stabilizes the interface. The only other interdomain interaction observed is a hydrophobic stacking of Trp413 and Lys575. IgIV and IgV of the IgV–Fn3I structure have fold and interdomain link-region interactions that are very similar to those observed in the IgV–Fn3II structure. However, the IgV–Fn3II link is significantly different in the two structures. The presence of a HEPES molecule appears to play an important role in stabilizing the conformation between IgV and Fn3I in this crystal structure (see Figure S2 available online). The HEPES molecule makes hydrophobic contacts with Tyr411, Ala493, Asp494, Val495, Pro496, Ser497, Ser498, and Val525, and the complex is further stabilized by hydrogen bonds through two water molecules. The observed tilt angle between Fn3I and Fn3II is 57°. When examining the linkage between Fn3I and Fn3II, no hydrogen bonds or other types of interactions are evident.

Molecular dynamics simulations were conducted on a number of recombinant OCAM proteins. Ten-nanosecond simulations on monomeric IgI–III domains showed several large folding and twisting motions around the hinges between the domains. The movements around the IgI–II hinge were significantly larger than the movements around the IgII–III hinge, and the tilt angles of the domains varied from 12° to 73° around the IgI–III hinge (compared with only 28°–44° for the IgII–III hinge) (Table S1). Two simulations of IgIV–Fn3I with and without the HEPES molecule showed significant differences. The simulation with the HEPES molecule bound to the protein only showed small variations compared with the crystal structure, and the simulation without HEPES displayed significant movement of the domains. During this simulation the tilt angle between the IgV and Fn3I domains changed from 98° (similar to the crystal structure) toward an equilibrium value of 65° (which is close to the tilt angle observed in the IgIV–Fn3II structure). The amino acids Leu395, Asp396, Leu490, Ala491, and Leu492 have been identified (by analysis backbone torsional angle changes) as being primarily responsible for the change in the tilt angle. The difference between the two simulations of IgIV–Fn3I with and without the HEPES molecule indicates flexibility in the IgV–Fn3I linker region and that the HEPES molecule is stabilizing the conformation observed in the crystal structure.

The simulations of the IgIV–Fn3II structure (four domains and three hinges) showed some flexibility of the hinges, with the Fn3II–Fn3I hinge being most flexible. Thus, the tilt angle between IgV and IgV varied from 24° to 27°, between IgV and Fn3I from 50° to 52°, and between Fn3I and Fn3II from 57° to 82°. During the simulations, no conformations were even remotely similar to the IgV–Fn3I structure with a HEPES molecule bound. Thus, analysis of the longest recombinant OCAM proteins indicated that most of the flexibility of the OCAM ectodomain originates from the IgI–II hinge.

which has been shown to be implicated in polysialylation of the NCAM1 IgV domain (Colley, 2010; Mendiratta et al., 2006), is not present in OCAM, and indeed OCAM has been reported not to be polysialylated (Yoshihara et al., 1997).

**Interdomain Arrangements**

The link region between OCAM IgI and IgI appears almost linear, with tilt and twist angles of 4°–12° and 144°–149°, respectively (Figure 2A). The most prominent interdomain interaction observed is a hydrophobic stacking of Arg193 and Tyr112. Arg193 also forms hydrogen bonds to the carbonyl oxygen atoms of residues Gln113 and Ser141. Altogether, this arrangement appears to confer rigidity to the linking region between the two domains. A similar stacking interaction is found in the NCAM1 structure, where the corresponding arginine forms an interdomain salt bridge to Glu30 (Uniprot code P13596). With a tilt angle of 28°–31°, IgII and IgIII are positioned in a near-linear arrangement. The conformation is stabilized by interdomain hydrogen bonds and salt bridges. Glu125 forms a salt bridge to Lys286, and the side chain of Lys127 forms a hydrogen bond to the carbonyl oxygen atom of Ala287. The tilt angle in OCAM IgI–II is very similar to that of NCAM1. A salt bridge between OCAM residues Glu125 and Lys127 is also seen in NCAM1 (Glu128 and Lys293). The equivalent of the hydrogen bond between OCAM residues Lys127 and Ala287 is also formed in NCAM1 between Lys130 and Ala294. The OCAM IgII G strand and the IgIV A strand form one long strand in the structure. Along with the interdomain hydrogen bond formed between residues Arg224 and Gln301, this long strand may stabilize the domain interface at a twist angle of 24°. Similarly, the OCAM IgIV G strand is extended into the IgV domain, where it forms one long β strand together with the A strand of IgV. Many other interdomain interactions between IgIV and IgV are present, including hydrogen bonds between Asp396 and Asn427, Glu398 and His480, and Glu398 and Asn427. The observed tilt angle is 27°–28°, again corresponding to a near-linear arrangement of the domains.

The interdomain link between IgV and Fn3I in the OCAM IgV–Fn3I structure is more bent than the linker regions between Ig domains, with a tilt angle of 52°. One interdomain hydrogen bond between Trp413 and Gly524 stabilizes the interface. The only other interdomain interaction observed is a hydrophobic stacking of Trp413 and Lys575. IgIV and IgV of the IgV–Fn3I structure have fold and interdomain link-region interactions that are very similar to those observed in the IgV–Fn3II structure. However, the IgV–Fn3II link is significantly different in the two structures. The presence of a HEPES molecule appears to play an important role in stabilizing the conformation between IgV and Fn3I in this crystal structure (see Figure S2 available online). The HEPES molecule makes hydrophobic contacts with Tyr411, Ala493, Asp494, Val495, Pro496, Ser497, Ser498, and Val525, and the complex is further stabilized by hydrogen bonds through two water molecules. The observed tilt angle between Fn3I and Fn3II is 57°. When examining the linkage between Fn3I and Fn3II, no hydrogen bonds or other types of interactions are evident.

Molecular dynamics simulations were conducted on a number of recombinant OCAM proteins. Ten-nanosecond simulations on monomeric IgI–III domains showed several large folding and twisting motions around the hinges between the domains. The movements around the IgI–II hinge were significantly larger than the movements around the IgII–III hinge, and the tilt angles of the domains varied from 12° to 73° around the IgI–III hinge (compared with only 28°–44° for the IgII–III hinge) (Table S1). Two simulations of IgIV–Fn3I with and without the HEPES molecule showed significant differences. The simulation with the HEPES molecule bound to the protein only showed small variations compared with the crystal structure, and the simulation without HEPES displayed significant movement of the domains. During this simulation the tilt angle between the IgV and Fn3I domains changed from 98° (similar to the crystal structure) toward an equilibrium value of 65° (which is close to the tilt angle observed in the IgIV–Fn3II structure). The amino acids Leu395, Asp396, Leu490, Ala491, and Leu492 have been identified (by analysis backbone torsional angle changes) as being primarily responsible for the change in the tilt angle. The difference between the two simulations of IgIV–Fn3I with and without the HEPES molecule indicates flexibility in the IgV–Fn3I linker region and that the HEPES molecule is stabilizing the conformation observed in the crystal structure.

The simulations of the IgIV–Fn3II structure (four domains and three hinges) showed some flexibility of the hinges, with the Fn3II–Fn3I hinge being most flexible. Thus, the tilt angle between IgV and IgV varied from 24° to 27°, between IgV and Fn3I from 50° to 52°, and between Fn3I and Fn3II from 57° to 82°. During the simulations, no conformations were even remotely similar to the IgV–Fn3I structure with a HEPES molecule bound. Thus, analysis of the longest recombinant OCAM proteins indicated that most of the flexibility of the OCAM ectodomain originates from the IgI–II hinge.
Dimerization of NCAM 2 IgI–IgII

The IgI–II and IgI–III structures reveal one significant protein–protein interface with an area of ~980 Å² in the IgI–II structure with ten hydrogen bonds and six salt bridges in the interface (residues involved in dimerization can be seen in Figure 1). In the IgI–III structure the interface area is only 620 Å², possibly due to a smaller angle between the molecules constituting the dimer than in the IgI–II dimer (113° versus 133°) (Figure 3A). The dimer interfaces obtained from the IgI–II and IgI–III structures, respectively, are shown in Figure 3B. Approximately 80% of the residues involved in this dimer interface are conserved between NCAM1 and OCAM (Figure 1), and the OCAM interface is very similar to that observed in the NCAM1 IgI–II and IgI–III structures (Colley, 2010; Kasper et al., 2000; Sorkina et al., 2003). Analysis of the crystal packing interactions in the four other structures did not reveal any other significant interfaces.

To investigate the IgI–IgII dimerization in solution, gel filtration chromatography was applied. The protein eluted as one peak at an elution time corresponding to the dimeric form of the protein (Figure S3A). Moreover, OCAM IgI–III, comprising 291 amino acids, eluted as a protein with a higher molecular weight than OCAM IgIV–Fn3I, comprising 400 amino acids (Figure S3A). Moreover, OCAM IgI–III, comprising 291 amino acids, eluted as a protein with a higher molecular weight than OCAM IgIV–Fn3I, comprising 400 amino acids (Figure S3B), indicating that IgI–III dimerizes in solution. To further investigate the structure of OCAM IgI–III in solution, the protein was subjected to small-angle X-ray scattering (SAXS) analysis. The molecular weight of the solute was estimated to be 72 kDa, corresponding to the expected molecular weight of a dimer (69 kDa). The pairwise distance distribution function (P(r)) derived from the solution-scattering curve (Figure S4) indicates a maximal dimension of the protein dimer of 150 Å. This dimension is significantly larger than the maximal distances observed within the crystal structure of the IgI–III dimer (~130 Å). The modular buildup of the dimer is evident from the P(r), which

Table 1. Data Collection and Processing Statistics for OCAM IgI–II, IgI–III, IgI–IV, IgI–Fn3I, and IgI–Fn3II

<table>
<thead>
<tr>
<th>Construct</th>
<th>OCAM IgI–II</th>
<th>OCAM IgI–III</th>
<th>OCAM IgI–IV</th>
<th>OCAM IgI–Fn3I</th>
<th>OCAM IgI–Fn3II</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray source</td>
<td>I911-2, MAX-Lab, Lund, Sweden</td>
<td>ID23-1, ESRF, Grenoble, France</td>
<td>ID29, ESRF, Grenoble, France</td>
<td>ID23-2, ESRF, Grenoble, France</td>
<td>I911-5, MAX-Lab, Lund, Sweden</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.041</td>
<td>0.979</td>
<td>0.976</td>
<td>0.873</td>
<td>0.908</td>
</tr>
<tr>
<td>Space group</td>
<td>P₃2₁</td>
<td>P₄₂₁</td>
<td>P₂₁</td>
<td>P₂₁</td>
<td>P₄₂₁</td>
</tr>
<tr>
<td>Unit-cell parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a (Å)</td>
<td>42.06</td>
<td>114.95</td>
<td>38.39</td>
<td>43.81</td>
<td>139.66</td>
</tr>
<tr>
<td>b (Å)</td>
<td>42.06</td>
<td>114.95</td>
<td>106.78</td>
<td>45.81</td>
<td>139.66</td>
</tr>
<tr>
<td>c (Å)</td>
<td>202.91</td>
<td>46.00</td>
<td>188.74</td>
<td>45.72</td>
<td>47.79</td>
</tr>
<tr>
<td>ß (°)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>103.58</td>
<td>–</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>28.99–1.82</td>
<td>20.0–2.00</td>
<td>24.57–3.00</td>
<td>20.2–1.98</td>
<td>62.46–2.51</td>
</tr>
<tr>
<td>Number of unique</td>
<td>19,558</td>
<td>21,021</td>
<td>15,703</td>
<td>12018</td>
<td>16752</td>
</tr>
<tr>
<td>observations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.6 (4.5)</td>
<td>5.9 (6.1)</td>
<td>4.7 (4.7)</td>
<td>3.7 (3.6)</td>
<td>7.1 (7.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.5 (95.9)</td>
<td>98.7 (98.7)</td>
<td>91.9 (96.1)</td>
<td>96.8 (90.9)</td>
<td>99.9 (100)</td>
</tr>
<tr>
<td>l/d(å)</td>
<td>23.1 (3.9)</td>
<td>15.60 (2.80)</td>
<td>16.60 (3.80)</td>
<td>10.40 (2.6)</td>
<td>14.60(2.90)</td>
</tr>
<tr>
<td>Rsym (%)</td>
<td>3.5 (50.4)</td>
<td>9 (25)</td>
<td>6 (39)</td>
<td>8.5 (58.9)</td>
<td>10.1 (72.6)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rwork (%)</td>
<td>20.3</td>
<td>18.2</td>
<td>21.3</td>
<td>20.0</td>
<td>19.4</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>24.0</td>
<td>23.4</td>
<td>28.7</td>
<td>26.6</td>
<td>24.1</td>
</tr>
<tr>
<td>Total atoms</td>
<td>1,696</td>
<td>1,730</td>
<td>4,283</td>
<td>1,647</td>
<td>2,594</td>
</tr>
<tr>
<td>Protein</td>
<td>1,501</td>
<td>1,458</td>
<td>4,209</td>
<td>1,526</td>
<td>2,339</td>
</tr>
<tr>
<td>Water</td>
<td>175</td>
<td>210</td>
<td>31</td>
<td>93</td>
<td>183</td>
</tr>
<tr>
<td>Ligands</td>
<td>20</td>
<td>62</td>
<td>43</td>
<td>28</td>
<td>72</td>
</tr>
<tr>
<td>Rmsd bond lengths (Å)</td>
<td>0.007</td>
<td>0.005</td>
<td>0.006</td>
<td>0.007</td>
<td>0.008</td>
</tr>
<tr>
<td>Rmsd bond angles (°)</td>
<td>1.052</td>
<td>0.732</td>
<td>0.978</td>
<td>1.149</td>
<td>1.157</td>
</tr>
<tr>
<td>Mean B (Å²)</td>
<td>35.3</td>
<td>26.5</td>
<td>96.58</td>
<td>39.7</td>
<td>51.4</td>
</tr>
<tr>
<td>Ramachandran plot (%)</td>
<td>92.4; 6.4; 1.2; 0.0</td>
<td>90.0; 10.0; 0.0; 0.0</td>
<td>76.3; 20.8; 2.1; 0.8</td>
<td>91.5; 7.9; 0.6; 0.0</td>
<td>87.4; 11.9; 0.4; 0.4</td>
</tr>
</tbody>
</table>

Values in parentheses are for the highest resolution shell.  

a $R_{\text{sym}} = \frac{\sum_i |I_{hkl,i} - <I_{hkl}>|}{\sum_i |<I_{hkl}>|}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices $h,k,l$ and $<I_{hkl}>$ is the mean intensity of that reflection.  

b $R_{\text{work}} = \frac{\sum_i |I_{hkl,i} - F_{o,hkl}|}{\sum_i |F_{o,hkl}|}$, where $|F_{o,hkl}|$ and $|F_{c,hkl}|$ are the observed and calculated structure factor amplitudes.  

c $R_{\text{free}}$ is equivalent to $R_{\text{work}}$ but calculated with reflections omitted from the refinement process (5% of reflections omitted).  

d Percentage of total number of residues in the “favoured,” “allowed,” “additionally allowed,” and “disallowed” regions, respectively, of the Ramachandran plot according to PROCHECK definitions (Laskowski et al., 1993).
Figure 5. Structural Model of the Entire Ectodomain of OCAM

Structural Model of the Entire Ectodomain of OCAM

Structure

Figure 5. Structural Model of the Entire OCAM Ectodomain Built from the Overlapping IgI–II, IgII–III, IgIII–IV, and IgIV–Fn3II Structures

On the model, N-glycosylations are modeled as the complex sialylated fucosylated triantennary type. The glycosylation corresponding to the position of the polysialylation in NCAM is colored orange. The rest of the glycosylations are colored red. See also Figures S1 and S3.

clearly shows the presence of a typical distance around 20 Å, corresponding to the radius of gyration (Rg) of an Ig domain. An ab initio bead model can be fitted by the IgI–IgIII crystal structure in which the intermolecular angle (angle between molecules in the dimer) is opened an additional 28°.

Structural analysis of the dimerization site indicated that Leu27, Phe39, and Arg77 contribute significantly to the interaction interface. To verify the role of these amino acids, the following two double mutants were produced: Leu27Ser + Phe39Ser (IgI–ILF) and Phe39Ser + Arg77Glu (IgI–IFR). Both proteins eluted as broad peaks, corresponding to combined monomeric and dimeric forms of the proteins, whereas the intact protein eluted as a single peak at elution time corresponding to the dimeric form (Figure S3A). Therefore, the introduced mutations partly disrupt the IgI–IgII dimer interface.

Structural Model of the Ectodomain

A model of the entire extracellular part of OCAM was built from overlapping IgI–II, IgII–III, IgIII–IV, and IgIV–Fn3II structures (Figure 5). The model represents a complete dimer of the OCAM ectodomain interpreted as a trans-interacting homodimer with a cell-to-cell distance of about 20 nm.

DISCUSSION

To understand the structural context for homophilic binding of Ig CAMs, we previously attempted to crystallize NCAM1 and L1 ectodomain fragments consisting of more than four domains, but these crystals only diffracted to low resolution (Kulahin et al., 2004, 2005a). In the present study we chose another strategy and crystallized several overlapping constructs of the OCAM ectodomain. This approach allowed the construction of a structural model of the entire ectodomain based on six overlapping structures of truncated versions of the protein. Altogether, the high consistency in tilt and twist angles (Figure 2A) in combination with the results of the molecular dynamics simulations support the notion that the average conformations of the various recombinant OCAM proteins resemble the conformations seen in the crystals. Consequently, a structural model of the entire OCAM ectodomain was assembled from the crystallized overlapping modular constructs. The model is shown in Figure 5. It consists of an essentially linear segment composed of the IgI–V domains, followed by bent hinges between the IgV and Fn3I domains and between the Fn3I and Fn3II domains. The tilt and twist angles in IgI–II and IgI–II are comparable to the angles observed in NCAM1 (Figure 2B) (Kasper et al., 2000; Soroka et al., 2003). Furthermore, the angles between IgI and IgII are also within the same range as the angles observed in the crystal structures of VCAM IgI–II (Wang et al., 1995) and the EC domains in N- and E-cadherin (Koch et al., 1999). These findings further support a model in which the entire OCAM IgI–V segment can adopt an essentially linear conformation, with variations of only 10°–15° in the hinges.

The bend between OCAM IgV and Fn3I in the IgIV–Fn3II structure is caused by a HEPES molecule bound in the hinge region, and the structure can represent one of several OCAM conformations. This is consistent with findings from electron microscopy that identified a hinge region around IgV in NCAM1 (Hall and Rutishauser, 1987). The bend between OCAM IgV and Fn3I is less pronounced in the OCAM IgIV–Fn3II structure. However, models of the entire ectodomain based on the IgIV–Fn3II structure (with a tilt angle of 98°) and the IgIV–Fn3II structure (with a tilt angle of 52°) result in approximately the same intercellular distances, and both models imply that the Ig domains can be positioned parallel with and in quite close proximity to the cell surface. The residues that interact with the HEPES molecule are conserved between NCAMs from several species (Figure S2) and may, therefore, form a regulatory binding site that is functionally important for all NCAMs.

The analysis of the IgIV–Fn3II structure did not reveal any interdomain interactions between the Fn3 domains, which together with the results from the molecular dynamics simulations indicate that the hinge region between these domains is one of the flexible regions in the OCAM ectodomain. This interpretation is supported by the structure of NCAM1 Fn3I–II, which also indicates that the link between the NCAM1 Fn3I and Fn3II is flexible (Carafoli et al., 2008).

Protein-protein interactions in crystals may represent biologically important interactions, and several publications have shown that biologically important adhesion interactions can be found in the crystalline phase, although they are not detected in solution (Freigang et al., 2000). A recent structure model of OCAM IgI suggested that this module exhibits domain swapping created by an interchange of the N-terminal A1 β strands (Rasmussen et al., 2008). However, in the present study, neither the IgI–II nor the IgI–III structure demonstrated any domain swapping between the IgI domains, and in the six OCAM structures presented in this study, the IgI–II dimerization was the only homophilic interaction observed in more than one crystal structure. This interface is composed of essentially three linear epitopes, formed by residues equivalent to residues forming the IgI–II dimer in NCAM1 (Figure 1). Five interface residues (Lys38, Phe39, Thr79, Tyr81, and Arg189) are conserved between NCAM1 and OCAM, suggesting that they play a primary
molecules comprising the NCAM1 and OCAM dimers (Figure 3A) each other (Yoshihara et al., 1997). The angles between the sequence may explain why NCAM1 and OCAM do not bind Leu194 in NCAM1, and these differences in the protein Glu191 residues of OCAM may provide specificity for homophilic role as contributors to affinity. Conversely, the Thr41, Arg77, and Glu191 residues of OCAM may provide specificity for homophilic binding. These residues correspond to Leu40, Thr80, and Leu194 in NCAM1, and these differences in the protein sequence may explain why NCAM1 and OCAM do not bind each other (Yoshihara et al., 1997). The angles between the molecules comprising the NCAM1 and OCAM dimers (Figure 3A) suggest plasticity of the dimer interface. This is especially pronounced in OCAM, where the intermolecular angles derived from the IgI–II and IgI–III structures differ by 20°. The difference in the angle can probably be explained by different crystal packing of the proteins. Solution-scattering experiments suggest even more pronounced plasticity of the dimer interface (Figure S4). Molecular dynamics simulations indicate that the IgI–II hinge is one of the most flexible ectodomain regions, further supporting the dependence of the intermolecular angle on the conformational restrictions on the protein (crystal versus solution conformation). The flexibility of the IgI–II hinge might be necessary for conformational adaptation during OCAM dimerization and is probably reduced after the dimer formation, resulting in a linear conformation of the region.

NCAM1-mediated homophilic cell adhesion has been the topic of several studies (reviewed in Soroka et al., 2010). The first models, based on cell or microsphere aggregation techniques, suggested that the IgI domain mediated the trans-homophilic interaction (Rao et al., 1992). Later, an IgI-IgII interaction was detected by surface plasmon resonance (Kiselyov et al., 1997), and this gave rise to a model in which trans-homophilic binding was mediated by IgI–II dimerization. The most recent model is based on the crystal structure of NCAM1 IgI–III (Soroka et al., 2003). Interestingly, this model interprets the IgI–IgII interaction as a cis-interaction, whereas homophilic NCAM1 trans-interactions are mediated by zipper-like interactions among IgI–IgIII, IgII–IgIII, and IgII–IgII (Figure S1). In contrast the results presented in this study clearly suggest that in OCAM the IgI–IgII interaction represents a trans-interaction (Figure 5), whereas no homophilic cis-interactions were identified.

In conclusion the presented structural model of OCAM suggests that the molecule forms trans-homophilic dimers through reciprocal IgI–IgII interactions, and that the Ig domains form a linear segment positioned parallel to the cell surface, followed by bends in the more membrane-proximal part of the molecule. Moreover, the majority of the flexibility of the structure seems to be restricted to the membrane-proximal hinges of the molecule, and to the IgI-IgII dimerization interface. The spatial arrangement may facilitate heterophilic interactions between the OCAM ectodomain and unidentified counter-receptors. Whereas little is known about OCAM interactions, NCAM1 is known to form numerous extracellular heterophilic interactions, including interactions between the NCAM1 Fn3 domains and the fibroblast growth factor receptor and interactions with several extracellular matrix molecules (Nielsen et al., 2010; Walmad et al., 2004).

The model of the OCAM ectodomain described here implies that homophilic OCAM interactions lead to an intercellular distance of about 20 nm. This intercellular distance is comparable to the 14 nm intercellular distance recently reported for the trans-heterophilic SIRPα-CD47 interaction (Hatherley et al., 2009). In support of these models, such distances between cell membranes are observed in chemical synapses (Hormuzdi et al., 2004).

**EXPERIMENTAL PROCEDURES**

**Protein Cloning and Expression** Constructs encoding recombinant protein were prepared using polymerase chain reaction amplification of human NCAM2 cDNA (Ensembl Gene ID ENSG00000154654; RZPD, Germany) for subcloning into the pCIPZz C plasmid (Invitrogen). All recombinant proteins consist of two N-terminal residues (Ser and Met) remaining from the cloning procedure, followed by amino acids from various recombinant human OCAM proteins (Swiss-Prot code O15394) and six C-terminal histidine residues. The IgI–II construct comprises amino acids 19–218, IgI–III amino acids 115–301, IgI–III amino acids 19–301, IgI–IV amino acids 209–400, IgI–Fn3II amino acids 302–593, and IgI–Fn3II amino acids 302–693. Two double mutants of IgI–II (IgI–IIcF and IgI–IIcF) were produced using a PhusionTM Site-Directed Mutagenesis Kit (Finnzymes). Mutations Leu72Ser and Phe39Ser and mutations Phe39Ser and Arg77Glu were introduced into the OCAM IgI–II and OCAM IgI–III mutants, respectively. All constructs were verified by DNA sequencing. The recombinant plasmids were linearized with SacI enzyme and used for transformation of the Pichia pastoris strain KM71H (Invitrogen). Transformation and selection were performed by the protocol supplied by the manufacturer. A preinduction culture was grown for 48 h in BMGH medium before transfer and continued growth and induction in BMNH medium for 24 h. The secreted proteins were purified using Ni-NTA (Qiagen) affinity chromatography. All proteins were deglycosylated using Endo Hf (New England Biolabs). As a final purification step, gel filtration chromatography was performed either in phosphate-buffered saline (PBS; Sigma; for SAXS experiments and analytical gel filtration chromatography) or 10 mM HEPES (Sigma) (pH 7.4), 20 mM NaCl (Sigma; for X-ray crystallography experiments). All protein constructs were concentrated to 3–8 mg/ml using Amicon Centrifugal Filter Units (Millipore).

**Gel Filtration Chromatography**

Gel filtration chromatography was performed with a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) connected to an Agilent FPLC system (GE Healthcare). The flow rate was 1 ml/min, and the temperature was 5°–6°C in all experiments. The column was equilibrated with PBS (pH 7.4) or 10 mM HEPES (pH 7.4), 20 mM NaCl. The column was calibrated in PBS (Figure S2C). The calibration set (GE Healthcare) included blue dextran 2000 (~2000 kDa), thyroglobulin (669 kDa), ferritin (444 kDa), aldolase (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa), and chymotrypsinogen A (25 kDa). Two milliliters of protein were injected onto the column.

**Crystallization and Crystal Structure Determinations**

Crystallization was achieved by hanging-drop experiments by mixing 1 μl protein and 1 μl reservoir solution (30% PEG-4000, 0.2 M sodium acetate trihydrate, 0.1 M tris hydrochloride (pH 8.5), for IgI–II; 30% PEG-4000, 0.2 M lithium sulfate, 0.1 M tris hydrochloride (pH 8.5), for IgI–II; 25% PEG-6000, 0.1 M HEPES (pH 7.0), for IgI–IV; 4% PEG-3350, 0.02 M calcium chloride, 0.1 M citric acid (pH 4.5), for IgI–III; 10% PEG-3350, 0.1 M ammonium phosphate, 0.01 M trimethylamine hydrochloride for IgIV–Fn3II; and 14% PEG-10000, 0.2 M calcium acetate, 0.1 M HEPES (pH 6.5), for IgIV–Fn3II). All chemicals were purchased from Sigma-Aldrich. Diffraction data on all six truncations of OCAM were collected at 100 K. Details of data collection and processing are presented in Table 1. The structures were determined by the molecular replacement method using the program Phaser (Storoni et al., 2004). Details on the models used for the molecular replacement method are provided in the Supplemental Experimental Procedures. Model building was performed using the program ARP/WARP (Perrakis et al., 1999) and manual correction using the program COOT (Emsley and Cowtan, 2004). The program PHENIX (Adams et al., 2002) was used for refinement, and the quality of the final structures was examined using PROCHECK (Laskowski et al., 1993). Analysis of binding interfaces was performed using PISA server (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html, accessed November 28, 2009).
Molecular Dynamics

Molecular dynamics simulations were performed with the Amber software suite, version 8 (Case et al., 2005). Parameters were from the Amber 9 force field (Wang et al., 2000) for all molecules, with the exception that the NAG and HEPS molecules in model IgV–Fn3I used the parameters from the gaff force field (Wang et al., 2004).

The complexes were solvated in octahedral boxes of water molecules, extending at least 8 Å outside the protein. Hydrogen atoms and water molecules were added using the LEAP module in Amber, assuming that all Asp and Glu residues were negatively charged and that the Lys and Arg residues were positively charged.

The solvent and hydrogen atoms were relaxed by a short molecular mechanics minimization (100 steps), followed by 20 ps molecular dynamics simulation. After relaxation, the entire system was equilibrated by 50 ps of constant pressure simulations, followed by 200 ps of constant volume simulations. Finally, a production simulation of 10 ns was performed for each system. All simulations were performed using a time step of 2 fs. Additional details are provided in the Supplemental Experimental Procedures.

Solution SAXS

SAXS data collection was performed at beamline I711 (MAXlab, Lund, Sweden) in the momentum transfer range 0.010 < q < 0.325 Å−1 (σ = 4σ sin(θ)/λ, where θ is half the scattering angle) using a wavelength of 1.05 Å and a MAR165 CCD detector. The sample exposure time was 10 min. Buffers were measured before and after sample exposure and averaged before background subtraction. Repeated exposure did not reveal any radiation damage. BioXAS RAW (Toft et al., 2008) software was used for radial averaging and background subtraction. The average molecular mass of the protein was estimated from the extrapolated forward scattering (I0) by using a reference solution of bovine serum albumin. The data were fitted to theoretical scattering curves of relevant structures using the program CRYSOL (Svergun et al., 1995). The indirect Fourier transform program GNOM (Svergun, 1992) calculated the P(r) and estimated the average Rg and the maximal distance (Dmax)

The complex (model IgV–Fn3I) was positively charged. Asp and Glu residues were negatively charged and that the Lys and Arg residues were positively charged.

ACKNOWLEDGMENTS

We thank Carlsbergfondet, Augustinus Fonden, Beckett-Fonden, and the Danish synchrotron user organization DANS.CATT for financial support. The authors declare that they have no conflict of interest.

REFERENCES


Kiseliov, V.V., Skladchikova, G., Hinaby, A.M., Jensen, P.H., Kulahin, N., Soroka, V., Pedersen, N., Tsetlin, V., Poulsen, F.M., Berezin, V., and Bock,