Structural analysis of human CEACAM1 oligomerization

Amit K. Gandhi1,5✉, Zhen-Yu J. Sun2,5, Yu-Hwa Huang1, Walter M. Kim1, Chao Yang1, Gregory A. Petsko3, Nicole Beauchemin4 & Richard S. Blumberg1

The human (h) CEACAM1 GFCC' face serves as a binding site for homophilic and heterophilic interactions with various microbial and host ligands. hCEACAM1 has also been observed to form oligomers and micro-clusters on the cell surface which are thought to regulate hCEACAM1-mediated signaling. However, the structural basis for hCEACAM1 higher-order oligomerization is currently unknown. To understand this, we report a hCEACAM1 IgV oligomer crystal structure which shows how GFCC' face-mediated homodimerization enables highly flexible ABED face interactions to arise. Structural modeling and nuclear magnetic resonance (NMR) studies predict that such oligomerization is not impeded by the presence of carbohydrate side-chain modifications. In addition, using UV spectroscopy and NMR studies, we show that oligomerization is further facilitated by the presence of a conserved metal ion (Zn++ or Ni+++) binding site on the G strand of the FG loop. Together these studies provide biophysical insights on how GFCC' and ABED face interactions together with metal ion binding may facilitate hCEACAM1 oligomerization beyond dimerization.
**Results**

Crystal structure of hCEACAM1 oligomer with two adhesion sites and simultaneous interactions through GFCC- and ABED face. To understand the structural basis for higher-order oligomerization and clustering of hCEACAM1, we solved the crystal structure of a tagless hCEACAM1 IgV domain at 2.2 Å resolution (Fig. 1a–c and Supplementary Figs. 1, 2). Crystallographic data
collection statistics and structural refinement are summarized in Table 1. The hCEACAM1 crystal structure revealed three hCEACAM1 molecules (a, b, c) in the asymmetric unit (Fig. 1a and Supplementary Fig. 1a–c) that exhibited an overall similarity in their structure based upon evidence of an anti-parallel beta-sandwich fold of all three molecules compared to a single hCEACAM1 molecule as previously described for a hCEACAM1 homodimer structure (PDB code 4QXW)\(^9\). We found that molecules (b) and (c) formed a dimer through GFCC’ face interactions (Supplementary Table 2), which resembled the WT dimer (PDB code 4QXW) with a C-alpha root mean square deviation (RMSD) of 0.66 Å (over 1489 atoms) and an interaction interface area of 856.2 Å\(^2\) (Fig. 1b), that was slightly larger than the interface area of a previously described homodimer (824.6 Å\(^2\), PDB code 4QXW)\(^9\). This GFCC’ face interaction of the oligomer was mediated by various hydrogen-bond interactions that included residues Phe(F)29, S32, Y34, V39, Gly(G)41, Gln(Q)44, Gln(Q)89, Ile(I)91, N97 and E99 and hydrophobic interactions involving F29, I91 and V39 residues as previously described for a hCEACAM1 GFCC’-mediated homodimer (Fig. 1c, Supplementary Figs. 1c, 2a, b and Supplementary Table 2).

We also observed changes in the GFCC’ interface in association with an oligomer relative to that involved in a homodimer. Specifically, we observed formation of two new hydrogen-bond interactions between residues E37 and Arg(R)38 of both molecules (b & c) in the oligomer that were not evident in the crystal structure of a hCEACAM1 dimer (PDB code 4QXW)\(^9\) (Supplementary Fig. 2c, d). In addition, compared to the asymmetric formation of seven hydrogen-bond interactions of the N97 residue in a hCEACAM1 dimer (PDB code 4QXW), we

### Table 1 Data collection and refinement statistics (molecular replacement).

| Data Collection | Human CEACAM1 Oligomer (PDB code 7RPP) |
|-----------------|----------------------------------------|
| Space group     | C121                                   |
| Cell dimensions | \[ \alpha, \beta, \gamma \text{ (Å)} \] 73.13, 86.06, 66.08, 90.0, 100.27, 90.0 |
| Resolution (Å)  | 29.86–2.20 (2.32–2.20)                 |
| \( R_{merge} \) (%) | 13.3 (76.7)                            |
| \( I/\sigma I \) | 7.3 (2.1)                              |
| Completeness (%) | 100 (100)                              |
| Redundancy      | 4.9 (4.7)                              |
| **Refinement**  |                                        |
| Resolution (Å)  | 29.86–2.20 (2.32–2.20)                 |
| No. reflections \(^a\) | 101,329 (13,993)                       |
| \( R_{work}/R_{free} \) | 18.4/23.7                             |
| No. atoms       | Protein 2520                           |
|                 | Ligand/ion 8                           |
|                 | Water 74                               |
| B-factors       | Protein 33.0                           |
|                 | Water 34.73                            |
| R.m.s.deviations| Bond-lengths (Å) 0.016                 |
|                 | Bond-angles (°) 1.78                   |

\(^a\)Values in the parentheses are for highest resolution shell.
observed that the N97 residues of molecule (b) and (c) in the oligomer participated in four symmetrical hydrogen-bond interactions in the formation of the GFCC’ interaction face (Supplementary Fig. 2a). This indicates that the GFCC’ face exhibits some degree of flexibility in accommodating the formation of a higher-order oligomer such that it’s surface area further increased relative to that observed with a homodimer (Supplementary Table 2).

Notably, molecule (b) showed an additional interaction site involving the ABED surface with an interface area of 482.1 Å² (Fig. 1b) which was mediated by interactions with molecule (a) through residues located at its respective ABED face (Fig. 1a–c and Supplementary Fig. 1a–b, 3a). This involved six hydrogen-bond interactions (Supplementary Table 3) between the residues Q26-I67 (using nomenclature convention here and after, where Q26 residue is from molecule (b) and I67 residue in italics is from molecule (a)), Q27-Q76, Q53-Pro(P)59, S93-G63 and N23-Q54 (Fig. 1c and Supplementary Figs. 1b, 3a). The residues Q26 of molecule (b) and I67 of molecule (a) mediated two hydrogen-bond interactions and formed the central site for the formation of the ABED face, wherein the side-chain atoms NE2 and OE1 of molecule (b) residue Q26 made two hydrogen-bonds of 3.0 Å and 3.1 Å with the main chain oxygen and nitrogen atoms of molecule (a) residue I67, respectively (Fig. 1c and Supplementary Figs. 1b, 3a). Consistent with the centrality of the Q26-I67 interaction in the formation of an ABED-based interface interactions, these two amino acids were the only contact sites observed in a previously described N97A mutant crystal structure (Supplementary Fig. 3b) which otherwise disabled hCEACAM1 IgV GFCC’ face-mediated interactions. In addition, molecule (b) residues N23 and Q27 made side-chain to side-chain hydrogen-bond interactions of 2.7 Å and 3.1 Å with molecule (a) residues Q54 and Q76, respectively (Fig. 1c and Supplementary Figs. 1b, 3a). Side-chain to main-chain backbone hydrogen-bonds of 3 Å between molecule (b) residue Q53 and molecule (a) residue P59 further strengthened the ABED face interactions (Supplementary Figs. 1b, 3a). In contrast to the GFCC’ interface, the ABED interface did not involve hydrophobic interactions.

Interestingly, molecule (b) residue S93, which is located within the FG loop, made a side-chain to backbone hydrogen-bond interaction of 2.8 Å through molecule (a) residue G63 (Supplementary Fig. 3a). This unique interaction between a GFCC’ face and ABED face residue demonstrates the coordination between CEACAM1 residues in the formation of oligomers. Taken together, this oligomer crystal structure revealed simultaneous interactions through a primary (GFCC’-GFCC’) and secondary (ABED-ABED) contact at an atomic resolution level thus demonstrating the presence of two adhesion sites on a single hCEACAM1 IgV molecule. Such observations provide a structural foundation for understanding hCEACAM1 oligomerization.

Human CEACAM1 GFCC’ face-mediates higher-order oligomerization and micro-cluster formation is supported by flexible ABED face interactions. Previous structural studies of wild-type and GFCC’ face mutants of the hCEACAM1 IgV domain and the determination of ligand binding to hCEACAM1 have underscored the importance of the GFCC’ face as representing a central force in hCEACAM1 interactions and monomer:dimer equilibrium. In addition, the N97A mutation which disables GFCC’-mediated dimerization (PDB code 6X01) supports the importance of the GFCC’ face in oligomerization as well as this mutant does not show proper ABED face-associations and only exhibits minor contacts via hydrogen-bond interactions involving residues Q26 and I67 (Supplementary Fig. 3b). Consistent with the central role of GFCC’ face in dimerization and subsequent oligomerization, we observed GFCC’ face mediated interactions between two oligomers that enabled the formation of higher-order oligomers (Fig. 4a), wherein molecule (a) of the oligomer present in the crystal asymmetric unit mediated GFCC’ face interactions with a symmetry mate of molecule (a) of the symmetry-related second oligomer (Fig. 4a). Superimposition of the oligomer present in the crystal asymmetric unit and symmetry-related second oligomer revealed the GFCC’ face

Glycosylation of residues N70, N77 and N81 sites are not predicted to impede GFCC’ and ABED face hCEACAM1 oligomer interactions. Recent NMR studies have aimed to understand the role of glycosylation in hCEACAM1 GFCC’ face-mediated dimerization with inconclusive results. To better understand the potential impact of glycosylation in the formation of hCEACAM1 GFCC’ and ABED face interactions, we modeled the N-linked sugars N-acetylglucosamine (NAG) and β-d-Mannose (BMA) on the N70, N77 and N81 residues within the ABED face residues of all three human CEACAM1 molecules (a, b, c) in the hCEACAM1 oligomeric crystal structure shown above (Fig. 2a–c). This manual modeling of the first 3 sugars of the native glycans was based on the mouse CEACAM1 crystal structure (PDB code 1L62) wherein a conserved N70 glycosylation site residue was resolved and shown to be linked to two molecules of NAG and a molecule of BMA (Supplementary Fig. 4a, b). Consistent with this, our previous crystal structure of a hCEACAM1 dimer (PDB code 4QXW) supported this modeling as it showed binding of octyl beta-D-glucopyranoside (BOG), a glycan mimic, near to residue N70 (Supplementary Fig. 4c) which overlapped with the N70-linked sugar moieties modeled from mouse CEACAM1 (Supplementary Fig. 4d). We thus optimized a glycosylated model of hCEACAM1 with N-linked sugars at residues N70, N77 and N81 of each CEACAM1 molecule of our oligomer structure. We observed that none of the modeled sugar molecules exhibited evidence that they caused steric-hindrance to the ABED face hydrogen-bond interactions (Fig. 2a–c). In addition, none of the sugar molecules affected or blocked any of the residues which mediate GFCC’ face interactions in the formation of a homodimer based upon this modeling (Fig. 2a–c). Further, as an independent measure of our manual modeling, sugar molecule modeling with the CHARMM glycosylation server showed a similar model of hCEACAM1 glycosylation without any steric clashes at the GFCC’ and ABED faces (Supplementary Fig. 5a). Consistent with these structural analyses, a previous study showed glycosylation doesn’t affect CEACAM1 GFCC’-face mediated interactions with a microbial ligand.

We also performed a nuclear magnetic resonance (NMR) spectroscopy study of 15N-labeled wild-type hCEACAM1 IgV protein in solution with or without BOG taking advantage of our previously described NMR assignments. This showed that addition of BOG caused no peak shift changes at the GFCC’-based dimer interface (Fig. 3a and Supplementary Fig. 6) and only caused a few minor peak shifts that were localized to residues Leu(L)18 and L73 that were not involved in ABED face interactions (Fig. 3b, c). This was consistent with the observed binding of BOG proximal to residues L18, N70 and L73 (Supplementary Fig. 4c) in a previously described WT (PDB code 4QXW) or I91A IgV mutant (PDB code 6XNT) crystal structure. Thus, NMR studies of hCEACAM1 and BOG showed no significant changes at the ABED and GFCC’ faces that would anticipate the formation of steric clashing due to glycosylation as predicted by the modeling of sugar molecules onto the hCEACAM1 crystal structure thus allowing for the formation of the oligomeric structure observed.
formed in both oligomers were mostly identical (Fig. 4b) with an (RMSD) of 0.6 Å (over 1536 atoms) and with similar GFCC'-face mediated interactions.

Our observation of a secondary interaction site in association with the ABED face as observed in the oligomeric structure provided an opportunity to better understand its role in the formation of higher-order oligomers. We therefore determined the crystallographic Debye-Waller factor (temperature factor or B factor) of all three molecules of the human CEACAM1 crystal structure reported here. This revealed higher thermal motion or dynamic mobility in the side-chains of some of the ABED face which is smaller in size with fewer potential interactions and thus serves as an auxiliary, flexible interface.

A conserved metal ion binding site in hCEACAM1 and bridging by Ni^{++} and Zn^{++}. One of the important structural features of the hCEACAM1 GFCC’ face is the involvement of the FG loop at the GFCC'-mediated dimeric interface, where residues such as Q89 of the F strand, E99 of the G strand and N97 within the FG loop form a hydrogen-bond network (Supplementary Fig. 1c). Of further importance to the structural integrity and conformation of the FG loop and G strand, we also observed that the main-chain nitrogen and carbonyl oxygen atoms of H105 make two hydrogen-bonds with the carbonyl oxygen of F9 and the main chain nitrogen atom of V11, respectively (Supplementary Fig. 8a). Moreover, our analysis of hCEACAM1 WT and V39A mutant structures (PDB code 2GK2, 6XNW) revealed evidence of hexadentate interactions between three H105 and three V106 hCEACAM1 residues contained within the G strands of the FG loop which coordinated

![Fig. 2 Glycosylated model of the hCEACAM1.](image-url)

Fig. 2 Glycosylated model of the hCEACAM1. a Modeling of the N-linked sugars N-acetylglucosamine (NAG) and β-d-Mannose (BMA) onto the N70, N77, and N81 residues of all three human CEACAM1 molecules observed in the crystal structure, wherein hCEACAM1 molecules (a in green, b in cyan, and c in magenta) are shown by ribbon diagram. Right inset shows structure of two N-acetylglucosamine (NAG) and one β-d-Mannose (BMA) by stick representation, which are modeled onto each set of N70, N77, and N81 residues in the crystal structure shown in left and abbreviated as NNB in the figure. For each sugar molecule, carbon atoms in orange, carbonyl oxygen in red and nitrogen in blue, are colored, respectively. Three sugars are shown for each hCEACAM1 molecule a, b and c (labelled in green, cyan and magenta, respectively). b Modeled N-linked sugars N-acetylglucosamine (NAG) and β-d-Mannose (BMA) onto the N70, N77, and N81 residues of human CEACAM1 molecules (a, b) do not appear to block any of the ABED face hydrogen-bond interactions shown. Modeled sugars are shown by stick diagrams and labeled as above. c Surface representation (green, cyan, and magenta) for molecules a, b, and c, respectively, and stick representation of modeled sugar molecules. GFCC’ and ABED face interactions are depicted by black oblong and three modeled NNB sugars on each hCEACAM1 molecule residues N70, N77 and N81 do not appear to block ABED and GFCC’ face interactions. NNB sugars and hCEACAM1 molecules are colored and labeled as above.
with Ni$^{++}$ in the hCEACAM1 structures (Supplementary Fig. 8b) or Zn$^{++}$ in the hCEACAM6 structure10 (Supplementary Fig. 8c). This resulted in the bridging of three hCEACAM1 or three hCEACAM6 molecules in the crystal structures, respectively. Interestingly, the ~90% sequence identity observed between the hCEACAM1 and hCEACAM6 IgV domains, (Supplementary Fig. 5b) which mediates their interactions21, includes conservation of the H105 and V106 residues7.

These studies suggest that these divalent cations may also contribute to the formation of hCEACAM1 oligomers through interactions in the G strand of the FG loop. We therefore pursued UV-spectroscopy studies of hCEACAM1 with various metal ions (Fig. 6a, b). In this assay, various metals were titrated at a 1:1 or 2.5:1 molar ratio relative to a 100 $\mu$M concentration of hCEACAM1. We found that Zn$^{++}$ or Ni$^{++}$, but not Ca$^{++}$ or Mn$^{++}$ or Li$^{+}$ ions, showed solution opacity changes and concentration-dependent aggregation of hCEACAM1 with greater induction observed upon addition of Zn$^{++}$ (Fig. 6a, b).

To confirm this, we next carried out NMR spectroscopy studies. $^{15}$N-labeled WT hCEACAM1 spectra showed loss of hCEACAM1 NMR peak intensities in the presence of varying concentrations of ZnCl$_2$ without the emergence of new peaks (Fig. 7a, b and Supplementary Fig. 9a–e). Further, we observed that the NMR samples turned from clear to translucent, to opaque or precipitated with increasing amounts of ZnCl$_2$ (from 20 $\mu$M to 250 $\mu$M) with complete reversal through chelation of the divalent cation through addition of 1 mM EDTA (Fig. 7c and Supplementary Fig. 9f). Similar solution opacity changes and loss of NMR spectra peaks were also observed for $^{15}$N-labeled WT hCEACAM1 in the presence of NiCl$_2$ (Fig. 7d, e). The NMR spectra peak loss caused by addition of Zn$^{++}$ or Ni$^{++}$ were nearly uniform for all residues as a result of CEACAM1 aggregation and/or precipitation (Supplementary Fig. 10). Further, addition of 100 mM imidazole, an analog for histidine side-chains, to $^{15}$N-labeled WT hCEACAM1 in the presence of 130 $\mu$M NiCl$_2$ reversed the reduction in the peak intensities of hCEACAM1 observed during binding to the Ni$^{++}$ ions (Supplementary Fig. 11). Consistent with these results, mutation of H105 to alanine abrogated Zn$^{++}$ or Ni$^{++}$-mediated aggregation of hCEACAM1 (Fig. 6c, d) and no cross-linking of Ni$^{++}$ was detected with hCEACAM1 under oxidative conditions30 (Supplementary Fig. 12). These structural analyses together with spectroscopy and NMR studies support the direct involvement of Zn$^{++}$ or Ni$^{++}$ binding to H105 residue in the G strand of the FG loop of the hCEACAM1 IgV domain in the formation of oligomers and micro-clusters.

Discussion

In this study, we investigated the mechanisms involved in the formation of hCEACAM1 multimers. Our atomic level structural studies show that a single hCEACAM1 IgV domain can simultaneously interact with two other hCEACAM1 IgV domains along the dominant GFCC’ and minor ABED face and thus helps to reconcile the presence of ABED (PDB code 2GK2) and GFCC’ (PDB code 4QXW) face-mediated interactions as observed in previously reported crystal structures8,9. This is interesting as previous surface plasmon resonance studies of rat CEACAM1 IgV predicted the existence of two binding sites on the IgV domain consistent with low-resolution topography studies of rat CEACAM1 which show evidence of CEACAM1 trimers and higher order oligomers23. Our studies thus provide confirmation for this and mechanistic insights into this observation based upon...
the atomic level resolution of a hCEACAM1 oligomer as reported here. Specifically, we observed that the GFCC'-based interface in this structure exhibited twice the area of the ABED interface. Consistent with this, the GFCC' interface involved 15 hydrogen-bond and 3 hydrophobic interactions as previously reported for a hCEACAM1 homodimeric IgV domain crystal structure (PDB code 4QXW)\(^9\), compared to 6 hydrogen-bond and no hydrophobic interactions for the ABED interface. The ABED interface in the hCEACAM1 IgV oligomer also included four glutamine residues (four-Q pocket) which were conserved among other hCEACAM family members. This conservation of glutamine residues at this interface is expected to promote flexibility in this portion of the molecule that included an increase in the size of its interface area and interactions between the R38 and E37 residues that have not been previously observed in a homodimer.

As the structural studies reported here and previously involved a non-glycosylated N-domain that might interfere with multimerization given their location within the IgV domain\(^8,26,27\), we also sought to understand the effects of hCEACAM1 glycosylation on the ABED and GFCC' face interactions. To do so, we modeled carbohydrate side-chain modifications and performed NMR spectroscopy that probed beta-octylglucoside (BOG) binding to hCEACAM1. This modeling and testing of our assumptions based upon the observed BOG interactions with hCEACAM1 predicted that the presence of sugar molecules at the hCEACAM1 glycosylation sites associated with N70, N77 and N81 would not interfere with the observed ABED and GFCC' face interactions. Further, B factor analysis of the hCEACAM1 oligomeric structure showed that the sites predicted to be associated with carbohydrate side-chain modifications (N70, N77, N81) as well as the amino acid residues along the ABED interface were characterized by high thermal motion and thus, flexibility. These results are further supported by a recent elegant study published at the time of this manuscript’s submission which directly demonstrates by NMR and structural (PDB code 7MU8) observations that dimer formation through hCEACAM1 GFCC' face interactions is observed in the presence of partially glycosylated hCEACAM1\(^31\).

Finally, we examined the role of metals on hCEACAM1 multimerization as previous studies showed Ni\(^{++}\) and Zn\(^{++}\) -mediated bridging of 3 hCEACAM1 molecules by coordination through histidine and valine residues in the G strands of the FG loops\(^26\). We noticed the presence of Zn\(^{++}\) coordination of the same G strand residues in a previously reported structure of hCEACAM6\(^10\). As such, we also investigated the biochemical role of Zn\(^{++}\) and Ni\(^{++}\) in regulating hCEACAM1 multimerization using UV spectroscopy, NMR and site-directed mutagenesis studies. This revealed evidence of reversible Zn\(^{++}\)-and Ni\(^{++}\)-mediated oligomerization of hCEACAM1 based upon competition with EDTA or imidazole, a histidine residue mimetic. Note that we didn’t observe any paramagnetic broadening of HSQC peaks with the addition of NiCl\(_2\). This suggest that the residual NMR peaks were from un-ligated hCEACAM1,
and the metal bound hCEACAM1 molecules were likely in aggregates or precipitates. Interestingly at the basal state, CEA-CAM1 is known to form cis-dimers via a GXXXG motif (G432–6436) embedded in the transmembrane domain and inside-out Ca\(^{++}\)-dependent calmodulin binding downregulates cis-homodimerization of hCEACAM1 into monomers; these further facilitate trans-homophilic binding and downstream signaling\(^{24,32}\). Our studies thus implicate an additional mechanism that regulates oligomerization which is associated with a distinct binding motif in the G strand of the FG-loop of the IgV domain and involves specific metals in a concentration-dependent process. Similar Zn\(^{++}\)- and Ni\(^{++}\)-mediated dynamic and reversible oligomerization in coordination through histidine residues was observed with APOBEC3G (A3G)\(^{33}\), DNA cytidine deaminase and with carbon monoxide dehydrogenase (CooJ)\(^{34,35}\), respectively, using similar experimental approaches as reported here. In these studies, NMR titration experiments showed 50 mM Zn\(^{++}\)-dependent A3G (300 µM) oligomerization and transmission electron microscopy (TEM) showed 1 mM Ni\(^{++}\)-dependent CooJ (100 µM) oligomerization, which were reversible in both cases in the presence of EDTA\(^{33,35}\).

Overall, our studies provide structural insights into how hCEACAM1 could exist as heterogeneous mixtures of oligomers and in diverse conformational states on the cell surface through primary GFCC\(^{'}\) and secondary ABED interfaces that are mediated by the IgV domain which we show is able to simultaneously bind two hCEACAM1 molecules (Fig. 8). In this model, homodimerization through the GFCC\(^{'}\) face enables highly flexible ABED-mediated interactions to form oligomers and higher-order oligomers. These are formed by symmetry mates of molecules a, b and c as observed in the crystal structure and are also based upon GFCC\(^{'}\) face interactions which may be further facilitated by metal ions such as Zn\(^{++}\) and Ni\(^{++}\) (Fig. 8). It is not known how this model relates to cis or trans interactions or their functional implications, however. It is equally plausible that it represents a means to direct hCEACAM1 to an inactive state on the cell surface when in cis or to intracellular signal transduction when the GFCC\(^{'}\) and consequently ABED interactions occur in trans across two cells. Nonetheless, our proposed model and the results of our biophysical and structural studies help to better understand previous observations and provide insights into the structural basis for the formation of hCEACAM1 oligomers (Supplementary Fig. 13) with implications for other hCEACAM family members. They may also potentially provide insights into the nature of the lower affinity heterophilic ligands for hCEACAM1 such as PD1\(^{13}\) and TIM-3\(^{39}\) which are likely to be highly influenced by the effects of avidity that would be facilitated by hCEACAM1 oligomerization.

**Fig. 6** UV-spectroscopy studies of hCEACAM1 IgV WT and H105A mutant with various metal ions. a) Titration of hCEACAM1 WT 100 µM (with buffer) and with various metals (Ni\(^{++}\), Zn\(^{++}\), Mn\(^{++}\), Li\(^{+}\), Ca\(^{++}\)) at 1:1 and 1:2.5 ratio. OD analysis at 340 nm revealed Zn\(^{++}\) or Ni\(^{++}\) caused concentration-dependent aggregation of the hCEACAM1 WT protein after 30-minute incubation. b) OD analysis of supernatant after centrifugation at 280 nm for WT protein and with various metals (Ni\(^{++}\), Zn\(^{++}\), Mn\(^{++}\), Li\(^{+}\), Ca\(^{++}\)) at 1:1 and 1:2.5 ratio. c) Titration of H105A mutant 100 µM (with buffer) and with various metals (Ni\(^{++}\), Zn\(^{++}\), Mn\(^{++}\), Li\(^{+}\), Ca\(^{++}\)) at 1:1 and 1:2.5 ratio. OD analysis at 340 nm revealed no Zn\(^{++}\) or Ni\(^{++}\) mediated aggregation observed for the H105A mutant protein. d) OD analysis of H105A mutant supernatant after centrifugation at 280 nm alone and with various metals (Ni\(^{++}\), Zn\(^{++}\), Mn\(^{++}\), Li\(^{+}\), Ca\(^{++}\)) at 1:1 and 1:2.5 ratio. The mean values with standard deviations are shown in bar graph with error bars of the triplicate samples. Source data are provided in Supplementary Data 1.
Methods

Protein expression, 15N labeling/culture preparation, refolding and purification. hCEACAM1 IgV domain protein expression and purification were done mostly following our previously published protocols7,9. For unlabeled hCEACAM1 IgV protein, competent E. coli BL21 (DE3) were transformed with a pET21d vector containing human CEACAM1 IgV gene insert. Transformants were grown in 1 L of LB media supplemented with 100 µg/mL of ampicillin and induced with 1 mM IPTG after reaching an OD600nm of 0.8 at 37 °C. Next, the cultures were grown for more hours before harvesting cell pellets by centrifugation. For 15N-isotopic labeling of human CEACAM1 IgV, M9 minimal media containing 42 mM Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3350. The grown at 25 °C via sitting drop method in a reservoir solution containing 0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3350. The diffraction data was collected at the National Synchrotron Light Source (NSLS, Upton, NY, USA) beamline X25 and processed through iMosflm36. The structure of the hCEACAM1 oligomer in the C121 space group was determined using molecular replacement method using polyalanine model of our published hCEACAM1 homodimer crystal structure (PDB code 4QXW) as a search model and molecular replacement method using polyalanine model of our published hCEACAM1 homodimer crystal structure (PDB code 4QXW) as a search model and refined after various iterative rounds of simultaneous model building to R free values of 18.4% and 23.7% using Refmac37 integrated to CCP4 suite38 and COOT39, respectively. Data collection and refinement statistics are listed in Table 1.

Next, refolded unlabeled and 15N-labeled hCEACAM1 protein were concentrated and dialyzed against 4 liter of 10 mM Tris dialysis buffer pH 8.0 for 24 h and the dialysis buffer was changed 3 times. Dialyzed protein was filtered and then purified using MonoQ ion exchange column (GE Healthcare Life Sciences) using NaCl gradient of 0 mM to 200 mM in 10 mM Tris dialysis buffer pH 8.0. Peak fractions containing hCEACAM1 protein was verified using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions and further loaded onto a HiPrep 16/60 Sephacyr S-200 HR column (GE Healthcare Life Sciences) for size-exclusion chromatography in a buffer containing 10 mM HEPES, 150 mM NaCl pH 7.4 The final purity was >95%, as judged by SDS-PAGE.

Crystallization and structure refinement of hCEACAM1 oligomer. Cell pellets containing expressed hCEACAM1 IgV domain was purified using our previously published protocols7,9. Purified protein was concentrated to 15 mg/ml and after the various rounds of the crystallization screening, diffraction quality crystals were grown at 25 °C via sitting drop method in a reservoir solution containing 0.2 M Ammonium sulfate, 0.1 M HEPES pH 7.5, 25% w/v Polyethylene glycol 3350. The diffraction data was collected at the National Synchrotron Light Source (NSLS, Upton, NY, USA) beamline X25 and processed through iMosflm36. The structure of the hCEACAM1 oligomer in the C121 space group was determined using molecular replacement method using polyalanine model of our published hCEACAM1 homodimer crystal structure (PDB code 4QXW) as a search model and refined after various iterative rounds of simultaneous model building to final R and Rfree values of 18.4% and 23.7% using Refmac37 integrated to CCP4 suite38 and COOT39, respectively. Data collection and refinement statistics are listed in Table 1.

All the figures and comparison of the crystallographic Debye-Waller factor (temperature factor or B factor) of all three molecules of the hCEACAM1 oligomer

![Fig. 7 15N-HSQC spectra of WT hCEACAM1 IgV domain with zinc chloride (ZnCl2) and with nickel chloride (NiCl2) and effect of EDTA. 15N-HSQC spectra of 190 µM 15N-labeled WT hCEACAM1 IgV domain alone (panel a), peaks in red), and in the presence of 250 µM ZnCl2 (panel b, weakened peaks in black), and both 250 µM ZnCl2 and 1 mM EDTA (panel c, recovered peaks in black). The weakened and missing backbone amide peaks caused by zinc bridging induced oligomerization/aggregation as shown in middle (panel d) are recovered after addition of EDTA (right panel c) to sequester zinc ions. 15N-HSQC spectra of 190 µM 15N-labeled WT hCEACAM1 IgV domain in the presence of 250 µM NiCl2 (panel d, weakened peaks in black), and both 250 µM NiCl2 and 1 mM EDTA (panel e, recovered peaks in black). The weakened and missing backbone amide peaks caused by zinc bridging as shown in middle (panel b) or nickel bridging as shown in middle (d) panel are recovered after addition of EDTA to sequester zinc or nickel ions.](image-url)
structure was done using PyMOL (DeLano Scientific). Sequence alignments of hCEACAM1 family members were done using Clustal Omega. The residues level hydrogen-bond interactions and interface area as observed in the hCEACAM1 oligomer crystal structure for GFCC and ABED face was determined using PDB PISA (proteins, interfaces, structures and assemblies).

Structural modelling and attachment of the N-linked sugars. Based on the superimposition of mouse CEACAM1 crystal structure (PDB code 1L6Z) onto to all three chains of hCEACAM1 oligomer using Pymol, wherein a conserved N70 glycosylation site residue was resolved in mouse CEACAM1 structure with two molecules of N- linked acetylgalcosamine (NAG) NAG and a molecule of β-d-Mannose (BMA), two molecules of NAG and one molecule BMA were modeled on the hCEACAM1 N70 residue first for all three molecules of the hCEACAM1 oligomer. Then based on N70-linked sugar molecules, attachment of the N-linked sugars was performed for N77 and N81 residues of all three molecules of the hCEACAM1 oligomer using Coot. Comparison of the N-linked sugars hCEACAM1 oligomer model and WT homodimer with bound BOG (PDB code 4QXW) was performed using Pymol. In addition to manual modeling, sugar molecule modeling was also performed with the CHARMM glycosylation server.

Mutagenesis, expression, and purification of H105A mutant. Human CEACAM1 H105A mutagenesis studies were carried out by using the PCR-based QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) with mutant forward primer sequence 5′-tctgagtaattaacgccaagtgcaggctgtctttcctatta-3′ and reverse primer sequence 5′-taatgaagaagcaaccggtcagtttgccgtgtattaactcgag-3′. The previously described pET21h plasmid containing human CEACAM1 IgV gene insert was used as the template for generating the H105A mutant. PCR reactions for introducing this H105A mutation were run for 16 cycles of 95 °C and 1 min at 35 °C, followed by 6 min at 68 °C. The resulting hCEACAM1-H105A mutant plasmid was verified by standard Sanger DNA sequencing. For expression and purification of H105A mutant, similar protocols were used as applied for WT expression and purifications.

UV-spectroscopy studies of hCEACAM1 WT and H105A with various metal ions. In this assay, 100 µM concentration of tagless hCEACAM1 WT or H105A purified protein in 10 mM HEPES, 150 mM NaCl, pH 7.4 was incubated with various metals (ZnCl2, NiCl2, MnCl2, LiCl, CaCl2) at a 1:1 or 2.5:1 molar ratio for 30 min. First, OD measurement was performed at 340 nm to assess the aggregation and then after centrifugation for 2 min at 10,000 rpm, clear supernatant OD was also measured at 280 nm. In addition, blank experiments with buffer (10 mM HEPES, 150 mM NaCl, pH 7.4) were performed with or without various metals at the highest concentration used in this study (250 µM). This blank experiment didn’t show any significant optical absorbance at 340 nm with all the metals.

Oxidative cross-linking studies of hCEACAM1 with nickel. Purified hCEACAM1 IgV (400 µM) was incubated with 250 µM NiCl2 for 30 min in HEPES buffer (10 mM HEPES, 150 mM NaCl, pH 7.4). Crosslinking was performed by the addition of KHSO4, (800 µM) for 30 min and then quenched by addition of equal volume of an SDS-PAGE denaturing buffer. Evaluation of crosslinking was performed by direct size assessment by SDS- PAGE electrophoresis.

NMR spectroscopy studies of hCEACAM1 with various metal ions. NMR samples of 190 µM 15N-labeled CEACAM1 were prepared with varying concentration of metal ion solutions in 20 mM HEPES, pH 7.4, 150 mM NaCl and 5% D2O. The 1H-NMR experiments were performed at 25 °C on a Bruker Avance II 600 MHz spectrometer equipped with a Prodigy Cryoprobe. The NMR data were acquired with 512 and 80 complex points in the direct HN and indirect 15N dimension respectively. NMR spectra were processed with Bruker Topspin software.

Statistics and reproducibility. The X-ray data and structure refinement statistics for the hCEACAM1 oligomer crystal structure is shown in Table 1. UV-spectroscopy studies were performed on triplicate samples (n = 3) for each metal titration experiment with hCEACAM1 as shown in in Fig. 6. The mean values with standard deviations are shown in bar graph with error bars.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The atomic coordinates and structure factors were deposited with RCSB accession code 7RPP. All relevant data are available upon request.
References

1. Gray-Owen, S. D. & Blumberg, R. S. CEACAM1: contact-dependent control of immunity. Nat. Rev. Immunol. 6, 433–446 (2006).
2. Kim, W. M., Huang, Y. H., Gandhi, A. & Blumberg, R. S. CEACAM1 structure and function in immunity and its therapeutic implications. Semin. Immunol. 42, 101296 (2019).
3. Kammerer, R. & Zimmermann, W. Coevolution of activating and inhibitory receptors within mammalian carcinoembryonic antigen families. BMC Biol. 8, 12 (2010).
4. Huang, J., Hardy, J. D., Sun, Y. & Shively, J. E. Essential role of biliary glycoprotein (CD66a) in morphogenesis of the human mammmary epithelial cell line MCF10A. J. Cell. Sci. 112, 4193–4205 (1999).
5. Ergün, S. et al. CEA-related cell adhesion molecule 1: A potent angiogenic factor and a major effector of vascular endothelial growth factor. Mol. Cell. 5, 311–320 (2000).
6. Beauchemin, N. et al. Redefined nomenclature for members of the carcinoembryonic antigen family. Exp. Cell Res. 252, 243–249 (1999).
7. Gandhi, A. K. et al. Structural basis of the dynamic human CEACAM1 monomer-dimer equilibrium. Commun. Biol. 4, 360 (2021).
8. Fedarovich, A., Tomberg, J., Nicholas, R. A. & Davies, C. Structure of the N-terminal domain of human CEACAM3: binding target of the opacity proteins during invasion of Neisseria meningitidis and N. gonorrhoeae. Acta Crystallogr. D. Biol. Crystallogr. 62, 971–979 (2006).
9. Huang, Y. H. et al. CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. Nature 517, 386–390 (2015).
10. Bonsor, D. A., Günther, S., Beadenkopf, R., Beckett, D. & Sundberg, E. J. Diverse oligomeric states of CEACAM IgV domains. Proc. Natl Acad. Sci. USA 112, 13561–13566 (2015).
11. Bonsor, D. A. et al. The Helicobacter pylori adhesin protein HopQ exploits the N-terminal domain interactions: structural analysis of the binding site. J. Biol. Chem. 294, 7601–7614 (2019).
12. Klaile, E. et al. Binding of Candida albicans to human CEACAM1 and mucin domain containing protein-3. Sci. Rep. 8, 17512 (2018).
13. Wojtowicz, W. M. et al. A human IgSF cell-surface interactor reveals a complex network of protein–protein interactions. Cell 182, 1027–1043 (2020).
14. Korotkova, N. et al. Binding of Dr adhesins of Escherichia coli to human T-cell immunoglobulin and mucin domain containing protein-3. J. Mol. Biol. 434, 433–448 (2015).
15. Königer, V. et al. Helicobacter pylori exploits human CEACAMs via HopQ for receptor in the CEA family. Embo. J. 21, 2076-2086 (2002).
16. Pinkert, J. et al. T cell-mediated elimination of cancer cells by blocking the carbon monoxide dehydrogenase accessory protein CooJ from Rhodospirillum rubrum. Inorganois 7, 84 (2019).
17. Shandilya, S. M. D. et al. Crystal structure of the APOBEC3G catalytic domain reveals potential oligomerization interfaces. Structure 18, 28–38 (2010).
18. Antony, R. et al. Nickel-induced oligomerization of the histidine-rich metallochaperone CooJ from Rhodospirillum rubrum. Inorganois 7, 84 (2019).
19. Alfano, M. et al. Nickel monoxide dehydrogenase accessory protein CooJ is a histidine-rich multidomain dimer containing an unexpected Ni(II)-binding site. J. Biol. Chem. 294, 7601–7614 (2019).
20. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. IMOSFML: A new graphical interface for diffusion-image processing with MOSFLM. Acta Crystallogr. D. Biol. Crystallogr. 67, 271–281 (2011).
21. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D. Biol. Crystallogr. 53, 240–255 (1997).
22. Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D. Biol. Crystallogr. 67, 235–242 (2011).
23. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797 (2007).

Acknowledgements

The authors acknowledge thankful support from all the NSLS beamline scientists (Upton, NY, USA) for X-ray data collection and thanks various NMR facilities including HMBS BIO-NMR (Boston, MA, USA), DFCI NMR core (Boston, MA, USA) facilities for NMR data collection. This work was supported by the NIH Grant R01DK051362-21 and the High Pointe Foundation to R.S.B. A.K.G. and Z.-Y.I.S. performed biophysical characterization, x-ray crystallography and NMR experiments. A.K.G. performed x-ray crystallization and structure determination. Z.-Y.I.S. carried out NMR experiments. Y.-H.H., W.M.K. and C.Y. performed analysis of hCEACAM1 oligomer. G.P. provided structural expertise in crystallization and structural comparison. R.S.B. established CEACAM1 project, and together with A.K.G., G.P., W.M.K., Z.-Y.I.S., Y.-H.H. and N.B. wrote the manuscript. R.S.B. and A.K.G. are the senior author of this paper.

Competing interests

The authors declare the following competing interests: R.S.B. has several issued and pending patents describing potential therapeutic strategies for regulating CEACAM1. The authors declare the following competing interests: R.S.B. has several issued and pending patents describing potential therapeutic strategies for regulating CEACAM1.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-022-03996-4.

Correspondence and requests for materials should be addressed to Amit K. Gandhi or Richard S. Blumberg.

Peer review information Communications Biology thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Ingrid Span and Gene Chong. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
