High resolution X-ray and NMR structural study of human T-cell immunoglobulin and mucin domain containing protein-3

Amit K. Gandhi1, Walter M. Kim1, Zhen-Yu J. Sun2,6, Yu-Hwa Huang1, Daniel A. Bonsor3, Eric J. Sundberg3,7,8, Yasuyuki Kondo1,9, Gerhard Wagner2, Vijay K. Kuchroo4, Gregory Petsko5 & Richard S. Blumberg1

T-cell immunoglobulin and mucin domain containing protein-3 (TIM-3) is an important immune regulator. Here, we describe a novel high resolution (1.7 Å) crystal structure of the human (h)TIM-3 N-terminal variable immunoglobulin (IgV) domain with bound calcium (Ca++) that was confirmed by nuclear magnetic resonance (NMR) spectroscopy. Significant conformational differences were observed in the B–C′, C′–C″ and C′–D loops of hTIM-3 compared to mouse (m)TIM-3, hTIM-1 and hTIM-4. Further, the conformation of the C–C′ loop of hTIM-3 was notably different from hTIM-4. Consistent with the known metal ion-dependent binding of phosphatidylserine (PtdSer) to mTIM-3 and mTIM-4, the NMR spectral analysis and crystal structure of Ca++-bound hTIM-3 revealed that residues in the hTIM-3 F–G loop coordinate binding to Ca++. In addition, we established a novel biochemical assay to define hTIM-3 functionality as determined by binding to human carcinoembryonic antigen cell adhesion molecule 1 (CEACAM1). These studies provide new insights useful for understanding and targeting hTIM-3.

T-cell immunoglobulin and mucin domain containing protein-3 (TIM-3), also known as Hepatitis A virus cellular receptor 2 (HAVCR2), is involved in regulating innate and adaptive immunity and has thus emerged as a target for numerous therapeutics that are under clinical development. There are three TIM family members in humans (TIM-1, TIM-3, and TIM-4) and eight members in mice (TIM-1 to TIM-8)1–11. TIM-3 is expressed on various cell types such as T-helper 1 (Th1) lymphocytes, dendritic cells and natural killer (NK) cells3,8,12,13. A major feature of TIM-3 is its potential role as a checkpoint inhibitor in immune responses to tumors, as well as its involvement in chronic viral infections3,7,14–16. Human (h)TIM-3 consists of a membrane distal variable immunoglobulin (IgV)-like domain involved in ligand interactions, a membrane proximal mucin domain and a transmembrane domain that is connected to an intracellular cytoplasmic tail involved in phosphotyrosine-dependent signaling3,9,17,18. Previous structural and biochemical studies have shown that mouse (m)TIM-3 and mTIM-4 bind phosphatidylserine (PtdSer) at the GFCC′ face of the IgV domain in a calcium (Ca++) dependent manner. PtdSer has also been implicated as a ligand for hTIM-310 although a crystal structure depicting this association has yet to be solved. The IgV-like domain of hTIM-3 has also been demonstrated to bind other ligands such as

1Division of Gastroenterology, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, 75 Francis Street, Boston, MA, 02115, USA. 2Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA, 02115, USA. 3Institute of Human Virology, School of Medicine, University of Maryland, 725 W Lombard St, Baltimore, MD, 21201, USA. 4Evergreen Center for Immunologic Diseases and Ann Romney Center for Neurologic Diseases, Harvard Medical School and Brigham and Women’s Hospital, 77 Avenue Louis Pasteur, Boston, MA, 02115, USA. 5Department of Neurology and Feil Family Brain and Mind Research Institute, Weill Cornell Medical College, New York, NY, 10021, USA. 6Present address: Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, 02215, USA. 7Department of Medicine, School of Medicine, University of Maryland, Baltimore, MD, 21201, USA. 8Department of Microbiology and Immunology, School of Medicine, University of Maryland, Baltimore, MD, 21201, USA. 9Present address: Division of Gastroenterology, Department of Internal Medicine, Graduate School of Medicine, Kobe University, Kobe, 650-0017, Japan. Amit K. Gandhi, Walter M. Kim and Zhen-Yu J. Sun contributed equally. Correspondence and requests for materials should be addressed to R.S.B. (email: rblumberg@bwh.harvard.edu)
carcinoembryonic antigen cell adhesion molecule 1 (CEACAM1)4, high mobility group protein B1 (HMGB1)20 and galectin-9,19,18. The study of these interactions suggests that hTIM-3 may exhibit overlapping but distinct structural characteristics that determine specificity for binding of these unique ligands that all coalesce on the IgV-domain. However, the specific conformational details of hTIM-3 have been too elusive to date to allow for a comprehensive understanding of these ligand interactions. Here we report the high resolution co-crystal structure of hTIM-3 IgV-domain with bound Ca\(^{2+}\) that provides a structural basis for understanding hTIM-3 at an atomic level.

## Results

### Overall structure and NMR validation of hTIM-3.

We first characterized the proper folding of purified recombinant hTIM-3 protein in its calcium bound state. Previous studies have demonstrated a melting temperature (T\(_m\)) for unbound hTIM-3 of approximately 45 °C10, which we confirmed by differential scanning fluorimetry. Titration of calcium increased the melting temperature consistent with increased stability of the Ca\(^{2+}\)-bound hTIM-3 complex (Supplementary Fig. 1a,b). In addition, we observed an excellent global fold structure of our purified Ca\(^{2+}\)-bound hTIM-3 in solution as judged by the large spread of assigned amide resonance peaks in the 15N-heteronuclear single quantum coherence (HSQC) spectrum of 15N/13C-labeled hTIM-3 protein, which consisted mainly of extended beta-strands (Fig. 1a). To obtain the structural details of hTIM-3 at an atomic level, we therefore determined the crystal structure of the Ca\(^{2+}\)-bound hTIM-3 IgV domain at 1.7 Å resolution (PDB code 6DHB, Table 1). This structure exhibited features common to the TIM family members2,9–11,21 by possessing a two anti-parallel \(\beta\)-sheet sandwich formed from front AGFCC′C″ and back BED faces, respectively, which were linked by the B-C, E-F, C″-D and A-B loops (Fig. 1b,c).

### Structural comparison with mTIM-3 and conformational flexibility of hTIM-3 B-C loop, C″-D loop and GF-CC′ cleft.

We next compared the hTIM-3 IgV domain crystal structure with that previously described for mTIM-3 (PDB code 2OYP)3. This revealed similarity in the overall structure with a root-mean-square deviation (R.M.S.D) of 0.59 Å (over 599 atoms) (Fig. 2a). For example, the AGFCC′C″ and BED face interactions that are important for stabilization of the IgV fold in mTIM-3 including inter-sheet interactions between Trp53 and Val93, a salt bridge between Arg81 and Asp104 and three disulfide bonds were all conserved in hTIM-3 (Supplementary Fig. 2a). This structure exhibited features common to the TIM family members2,9–11,21 by possessing a two anti-parallel \(\beta\)-sheet sandwich formed from front AGFCC′C″ and back BED faces, respectively, which were linked by the B-C, E-F, C″-D and A-B loops (Fig. 1b,c).
disulfide bonds rather than in a reduced state (average 28.4 $\pm$ 2.4 ppm) in solution. Similarly, we observed that 33 out of 107 residues (30.9%) were solvent accessible in the hTIM-3 IgV domain compared to 39 solvent accessible residues out of 108 (36.1%) in the mTIM-3 IgV domain.

Despite these similarities, there were significant structural differences between the mTIM-3 and hTIM-3 IgV domains that centered around the B-C, C-C′, C′-C″, and C″-D loops. Notably, the B-C loop of mTIM-3 possesses three proline residues (Pro 37, Pro43 and Pro50) whereas hTIM-3 has an additional proline residue at position 45 (Pro37, Pro42, Pro45 and Pro50). Together with the presence of an additional aromatic residue at position 39 (Phe39) (Fig. 2b), the B-C loop (Ala43-Asn47) of hTIM-3 exhibits substantial changes in its conformation signified by its transposition towards (~4 Å) the $\alpha$-helix connecting the D and E beta strands (Supplementary Fig. 2b). This involves a carbonyl oxygen atom of hTIM-3 residue Ala44 that makes hydrogen bonded interactions with the main chain amide nitrogen of Asn47. Another important structural difference observed between mTIM-3 and hTIM-3 is evidence for a shorter C″ strand present in hTIM-3 that could impact the conformation of the C″-D loop. In addition, unlike mTIM-3, residues Trp78 and Trp83 of hTIM-3 create a hydrophobic pocket that allows for hydrogen bond formation between the side chain of Thr79 and main chain amide nitrogen of Tyr82. Together, these differences in hTIM-3 result in no apparent outward bending of the C″-D loop in contrast to what is observed in the mTIM-3 structure (Supplementary Fig. 2c). These observed conformational differences in hTIM-3 are likely associated with flexibility in the B-C and C″-D loops regions, account for the missing assignments of the B-C and C″-D loop regions of the molecule in our NMR studies (Supplementary Table 1, Supplementary Fig. 3) and may provide a rationale for the significantly different B-C and C″-D loop conformations observed in a deposited hTIM-3 crystal structure that lacks bound Ca$^{++}$ (PDB code 5F71).

A characteristic feature of mTIM-3 and the other human TIM family members hTIM-1 and hTIM-4 is the formation of a unique cleft between the C-C′ and F-G loops in solution. This cleft represents a conserved binding site for PtdSer on mTIM-3 as well as hTIM-1 and hTIM-4. Although the cleft between the C-C′ and F-G loops of human and mouse TIM-3 is generally conserved, hTIM-3 features Val60 and Phe61 residues in the C-C′ loop in contrast to mTIM-3, which has Trp60 and Ser61 residues at these orthologous positions (Fig. 2b). Consequently, we observed a different cleft conformation in hTIM-3 compared to mTIM-3, with an extended distance between

| Data collection statistics | hTIM-3 with bound Ca$^{++}$ |
|---------------------------|----------------------------|
| Space Group               | P 212121                   |
| Cell constants            | 42.31 Å, 43.06 Å, 51.72 Å |
| a, b, c, $\alpha$, $\beta$, $\gamma$ | 90°, 90°, 90° |
| Resolution (Å)*           | 33–1.70 (1.73–1.70)        |
| No. of measurements*      | 98221 (2181)               |
| Unique reflections*       | 10823 (526)                |
| I/sigma I*                | 31.1 (10.5)                |
| Completeness (%)*         | 99.2 (90.5)                |
| Redundancy*               | 9.1 (4.1)                  |
| $R_{avey}$ (%)***         | 4.6 (8.5)                  |
| CC_{1/2}***               | 0.999 (0.989)              |

Table 1. Crystal information, data collection and refinement parameters. $R_{avey} = \sum |I - \langle I \rangle|/\sum \langle I \rangle$, where $I$ is the observed intensity and $\langle I \rangle$ is the weighted mean of the reflection intensity. $R_{avey} = \sum |F_o| - |F_c|/\sum |F_o|$, where $F_o$ and $F_c$ are the observed and calculated structure factor amplitudes, respectively. $R_{free}$ is the crystallographic $R_{avey}$ calculated with 5% of the data that were excluded from the structure refinement. Values in the parentheses are for highest resolution shell.
the C-C′ and F-G loops in mTIM-3 (~11.5 Å) compared to hTIM-3 (~10.0 Å). Further, the C-C′ loop aromatic residue Trp60 of mTIM-3, which has been shown to change the conformation of the mTIM-3 cleft to modulate PtdSer entry and binding, is not conserved in hTIM-3 but is instead replaced by a hydrophobic Val60 residue that lacks an aromatic side chain (Fig. 2b). These structural differences provide distinct cleft volumes of 152.3 Å³ for hTIM-3, as shown here, and 50.6 Å³ in presence of PtdSer. This highlights the flexibility of the GF-CC′ cleft volume to accommodate various potential ligands.

Human TIM family structural comparisons. We also compared the hTIM-3 crystal structure with other human TIM family members that share >25% sequence identity (Fig. 2c). We found that the IgV domain of hTIM-3 revealed significant structural similarity with hTIM-1 (yellow, PDB code 5DZO), and hTIM-4 (magenta, PDB code 5DZN) IgV domain crystal structures superimposed on each other. Loops are highlighted in italics. (e) C-C′ loop conformation differences between human TIM-4 (magenta), hTIM-3 (green) and hTIM-1(yellow). C-C′ (hTIM-1 and hTIM-3) and F-G (hTIM-1, hTIM-3 and hTIM-4) loops are labeled in italics for hTIM family members, whereas C-C′ loop (magenta) for hTIM-4 is labeled in italics and underlined. Residues Cys52 and Cys63 that form noncanonical disulfide bond are aligned in hTIM-3 and hTIM-1, whereas this disulfide bond between Cys52 and Cys63 is significantly deviated in hTIM-4 due to a different CC′ loop conformations and position of Cys63 in hTIM-4. Human TIM-4 CC′ loop and residue Cys63 are underlined. The disulfide bond between Cys52 and Cys63 is shown in orange for hTIM-3 and hTIM-1, and in magenta for hTIM-4.
residues Trp117 and Phe118 in hTIM-1 and hTIM-4 at these positions which are important for PtdSer binding and lipid bilayer penetration. This observation potentially explains the diminished biochemical affinity that hTIM-3 displays for PtdSer relative to hTIM-1 and hTIM-4. As noted above, the F-G loop of the TIM family members forms a cleft along with the C-C′ loop. It is therefore of interest that the polymorphic Arg111 residue of the hTIM-3 F strand, which is conserved in all human TIM family members, and Lys122 in the hTIM-3 G strand make important hydrogen bond interactions with main chain oxygen atoms of the C-C′ loop residues in hTIM-3.

Of further interest, we observed significant divergence in the C-C′ loop among the human TIM family members. Except for two cysteine residues, most of the C-C′ loop residues are not conserved among the human TIM family members (Fig. 2c). Despite these differences, the overall conformation of the hTIM-3 C-C′ loop is similar to that of hTIM-1. In contrast, we observed that there was a notable conformational difference between the C-C′ loop of hTIM-3 and hTIM-4 in which the C-C′ loop of the latter is observed to rotate inward (Fig. 2e). This is due to the deviation of a noncanonical disulfide bond located between Cys52 and Cys63 of hTIM-4 which confers a distinct C-C′ loop conformation. Consequently, this creates a much larger cleft volume in hTIM-4 (275.5 Å³) relative to either hTIM-3 (152.3 Å³) or hTIM-1 (87.69 Å³). Together, this suggests that the GFCC′ face of hTIM-3 is more similar to hTIM-1 than hTIM-4.

Determination of Ca²⁺ binding site of hTIM-3 by x-ray crystallography and NMR that coordinate ligand binding. The GF-CC′ cleft of mTIM-3, mTIM-4 and hTIM-4 have been defined to possess a metal ion-dependent ligand binding site that is chelated by Ca²⁺ and involved in interactions with PtdSer. Our previously reported NMR studies have demonstrated that the F-G loop of hTIM-3 exhibits large chemical shift index changes upon the addition of 10 mM Ca²⁺ (Fig. 3a,b), and that hTIM-3 binds to hCEACAM1 in the presence of Ca²⁺. As shown in Fig. 4a, we performed calcium titration experiments between 0 and 100 mM Ca²⁺ and determined a Ca²⁺ binding constant (Kd) of 27.2 mM for hTIM-3 by evaluating the peak shifts of the neighboring hTIM-3 C-C′ loop aromatic residue Phe61 through non-linear regression analysis (Fig. 4b,c). Our crystal structure further provides an atomic basis for this interaction and explains the binding mode of Ca²⁺ as a cofactor for hTIM-3 function. As observed for other TIM family members, the Ca²⁺ binding site of hTIM-3 is located in the F-G loop and involves residues Ile114, Gly116, Asn119, and Asp120 that coordinate a single Ca²⁺ cation (Fig. 4d,e, Supplementary Fig. 4), consistent with our NMR findings (Fig. 3a,b). Specifically, the main chain oxygen atoms of Ile114 and Gly116 and side chain oxygen atoms of Asn119 and Asp120 make equidistant (2.3 Å) bonds with the bound Ca²⁺ (Fig. 4d,e), leaving two additional sites for ligand-binding coordination. Residues Asn119 and Asp120 that make side chain interactions with Ca²⁺ are conserved across all human TIM family members (Fig. 2c). Interestingly, Ile114 is not conserved in hTIM-1, but the main chain oxygen atom of hTIM-1 residue His114 and hTIM-4 residue Val114 are expected to coordinate Ca²⁺ similarly to Ile114 in hTIM-3. It is further notable that the hTIM-3 residues Ile114, Asn119, and Asp120 exhibit a very low B-factor consistent with this region providing a highly conserved site in coordinating Ca²⁺ (Table 1).
Based on the structural similarity between the human and mouse TIM-3 IgV domains, we also modeled PtdSer at the Ca\(^{++}\) binding site of hTIM-3 using the reported co-crystal structure \(^2\) of PtdSer complexed to mTIM-3 (PDB code 3KAA) as a template. Our model predicts that the hTIM-3 F-G loop residues Asn119 and Asp120 interact with PtdSer through Ca\(^{++}\)-mediated binding \(^10,19\). Our model also predicts that the F-G loop residues Ile117 and Met118 in hTIM-3 do not form interactions with PtdSer in contrast to residues Trp117 and Phe118 found in hTIM-1 and hTIM-4. Thus, the Ca\(^{++}\)-bound hTIM-3 crystal structure confirms a conserved metal binding site observed for other TIM family members which is predicted to enable hTIM-3 binding to PtdSer described in previously biochemical studies \(^10\).

Biochemical assessment of hTIM-3 binding to hCEACAM1. The well-characterized biochemical assessment and structural studies of hTIM-3 as reported here also allowed us the opportunity to use our previous finding of hTIM-3 binding to hCEACAM1\(^4\) as a basis to provide evidence that the hTIM-3 protein studied here is functional. We therefore established an enzyme-linked immunosorbent assay (ELISA) assay to evaluate hCEACAM1-hTIM-3 interactions. First, we observed binding of a hTIM-3 Fc fusion protein in solution to plate-bound tagless hCEACAM1 IgV domain, as detected by an anti-human IgG Fc specific horseradish peroxidase conjugated antibody (Fig. 5a). Further, this interaction was blocked by a hTIM-3 C-C′ loop-derived peptide (amino acids 58–77) but unperturbed by a scrambled control peptide (Fig. 5b). Consistent with our previously reported hTIM-3 mutagenesis studies \(^4\), this hTIM-3 C-C′ loop-derived peptide possesses hTIM-3 residues Glu62 and Arg69 which are critical for hCEACAM1 binding. In an alternative approach, we applied the tagless hTIM-3 IgV domain used in crystallization experiments to a plate and observed binding of a glutathione-S-transferase (GST) tagged hCEACAM1 IgV-domain protein, but not GST alone (Fig. 5c). The latter interaction was blocked by a mouse anti-human CEACAM1 IgV-domain specific monoclonal antibody (5F4) \(^24\), but not by an isotype control antibody (Fig. 5d). Together, these studies show that the hTIM-3 IgV domain protein used in the structural studies exhibit functional binding activity.

Figure 4. Structural basis of hTIM-3 binding with Ca\(^{++}\). (a) Overlaid \(^{15}\)N-HSQC spectra showing chemical shift changes of hTIM3 backbone amides peaks induced by Ca\(^{++}\) binding, from 0 mM (red), 2 mM (orange), 5 mM (yellow), 10 mM (green), 20 mM (cyan), 50 mM (blue) to 100 mM (purple) concentrations. Selected residues of hTIM-3 GFCC′ face (Ile114 to Lys122) that show significant chemical shift changes are labeled. Some of peaks of these residues above 10 mM Ca\(^{++}\) concentration are not clear, possibly due to dynamic conformation exchange in solution. (b) \(^{15}\)N-HSQC peak shift monitoring of the hTIM-3 C-C′ and F-G loop aromatic residues Phe61 (F61) and Asn119 (N119) upon binding with 0–100 mM Ca\(^{++}\). (c) Determination of Ca\(^{++}\) binding constant (Kd) of 27.2 mM for hTIM-3 based on peak shifts of residue Phe61 determined by non-linear regression of the chemical shift change index relative to the Ca\(^{++}\) concentration. (d) Interactions of hTIM-3 F-G loop residues with bound Ca\(^{++}\) (red sphere) as observed in the crystal structure. Ile114, Gly116, Asn119, and Asp120 residues of hTIM-3 are highlighted in stick representation and these residues make equal distance coordination (2.3 Å) with Ca\(^{++}\). Residues Pro115 and Ile117 of hTIM-3 FG loop are also shown in a stick representation. (e) 2Fo-Fc map at 1.0 σ level of hTIM-3 residues Ile114, Pro115, Gly116, Asn119, and Asp120 with bound Ca\(^{++}\). Map is superimposed on the final refined model.
Discussion

We have performed a biophysical and structural assessment of the hTIM-3 IgV domain. To do so, we performed the first NMR analysis of a TIM family member and solved a high resolution (1.7 Å) crystal structure of hTIM-3 IgV domain. This has allowed us the opportunity to decipher the structural details of hTIM-3 at an atomic level. Since the description of the first crystal structure for mTIM-3 (PDB code 2OYP) in 2007, there is only one report of a low resolution (2.4 Å) hTIM-3 structure (Fig. 6a) described in the literature (PDB code 5F71)23. This crystal structure of hTIM-3 exhibited several structural features that are worthy of mention relative to the crystal structure described here. Notably, the previously reported crystal structure lacks a fully formed channel between the C-C′ and F-G loops of hTIM-3 structure (Fig. 6b), as observed in mTIM-3 9 and our hTIM-3 crystal structure reported here (Fig. 1c). In addition, this hTIM-3 crystal structure (PDB code 5F71) exhibits significant conformational differences in the B-C, C′-C″, and C″-D loops. (Fig. 6c). More specifically, this previously reported hTIM-3 structure completely lacks a C″ beta strand and instead contains an additional alpha helix with bound sodium (Na+). This additional helix with bound sodium was not observed in mTIM-3 9 and our hTIM-3 crystal structure. The finding of Na+ in PDB code 5F71 is perhaps consistent with the presence of sodium abundant crystallization conditions23. These overall structural characteristics are also associated with a high average B factor for all atoms (40 Å²)23, as judged by the PDB validation report (Fig. 6f). Most importantly, our structure, but not PDB code 5F71 23, is Ca++-bound to amino acids in the GFCC′ face. This is particularly important as numerous studies2,10,19 consider Ca++ to be an essential factor for TIM-family binding to its ligands suggesting our structure is representative of the biologically active state of hTIM-3. Thus, the hTIM-3 crystal structure reported here most likely resembles the in vivo biologically relevant Ca++-bound state of this important immunologically active molecule.

These studies also provide important insights into the mechanism by which the hTIM-3 GFCC′ face may bind its various ligands and highlights the similarities and conformational differences of hTIM-3 relative to other hTIM family members and mTIM-3. It is speculated that the presence of these significant structural differences between mTIM-3 and hTIM-3 may impact their species-specific functions. Moreover, using the hTIM-3 IgV protein used in the structural studies described, we established a novel ELISA assay to assess hTIM-3 functionality.

Figure 5. Human TIM-3 and human CEACAM1 binding studies by ELISA. (a) hTIM-3 Fc fusion protein but not hlgG-Fc fusion protein (0–4 μM) binds to tagless hCEACAM1 IgV domain with EC50 of 1.50 μM. (b) Blockade of the interaction between tagless hCEACAM1 and hTIM-3 Fc fusion protein with 10 μM of hTIM-3 C-C′ loop-derived peptide (amino acids 58–77) but not by 10 μM of scrambled peptide. (c) Binding of tagless hTIM-3 IgV domain with glutathione-S-transferase (GST) tagged hCEACAM1 IgV-domain protein (0–15 μM) with EC50 of 3.13 μM, but not with GST protein (0–15 μM). (d) Significant blockade of interaction between tagless hTIM-3 IgV domain and glutathione-S-transferase (GST) tagged hCEACAM1 IgV-domain fusion protein by a mouse anti-human CEACAM1 IgV-domain specific monoclonal 5F4 antibody (0–1 μM), but not by an isotype control MOPC antibody. (a–d) ELISA binding assays were performed in triplicate and the average values are shown with standard deviations. n.s., not significant. p values ≤*0.05, **0.01, and ***0.001.
through its ability to bind hCEACAM1 as previously shown. In addition, we showed that this interaction was blocked by a peptide derived from the hTIM-3 C-C' loop that contained amino acid residues previously shown to be involved in hCEACAM1 interactions and a mouse anti-human CEACAM1 antibody specific for the CEACAM1 N-domain. These latter studies further emphasize the importance of the GFCC' face of hTIM-3 in ligand binding and especially those that involve hCEACAM1.

Several polymorphisms in mouse and human TIM-3 have been previously identified and shown to be associated with alterations of ligand interactions and/or disease susceptibility. Our structural studies provide potential insights into these observations. For example, polymorphisms in the mTIM-3 B-C loop at the Pro43, Ser45 and Thr47 residues have been described that are associated with autoimmune and allergic disease susceptibility. Specifically, the allergy-susceptible BALB/c mouse strain possesses a Ser45 residue at this position in contrast to the DBA/2 mouse strain that includes Pro45 and exhibits resistance to allergy in models of airway hypersensitivity.

Interestingly, hTIM-3 exhibits an orthologous Pro45 residue at this position that results in an extra proline residing in the B-C loop of hTIM-3 compared to mTIM-3 and causes conformational changes in the B-C loop as described herein (Supplementary Fig. 2b). On the other hand, polymorphisms in hTIM-3 including Cys58Arg (rs201750016) and Arg111Trp (rs145478313), which would be predicted to disrupt the GFCC' cleft based upon our crystal structure, have been shown to affect the biochemical interactions between hTIM-3 and hCEACAM1.

The high resolution of our hTIM-3 crystal structure also allowed mapping of potential glycosylation sites on hTIM-3, which is of interest since TIM-3 interacts with galectin-9 through the glycans on the TIM-3 IgV domain. In comparison to the glycosylation sites of mTIM-3 that map to Thr44 (B-C loop), Asn74 (C'-C'' loop) and Asn100 (E-F loop) residues, the glycosylation sites of hTIM-3 are predicted to exist at residue Asn33 (A'-B loop), Asn100 (E-F loop) and Asn124 (G strand) (Fig. 7a-d). It is notable that while all three glycosylation sites for mTIM-3 are located at opposite side of the GFCC' face (Fig. 7a-c), the localization of the glycosylation sites of hTIM-3 were distinctly different. Specifically, the glycosylation site at residue Asn124 of hTIM-3 is present on the G strand which causes it to lie adjacent to the GF-CC' cleft and, interestingly, the glycosylation sites at Asn33 and Asn100 that create a continuous glycan surface on hTIM-3 where galectin-9 binds based on studies with mTIM-3.
Together, these differences may predict unique interactions of galectin-9 with hTIM-3 relative to mTIM-3 and the possibility that carbohydrate side-chain modifications at Asn124 may affect hTIM-3 interactions with ligands that bind to the GFCC′ face.

Recent studies have suggested that therapeutic agents capable of exhibiting optimal anti-tumor responses target hTIM-3 on the GFCC′ face where most known hTIM-3 ligands bind. Our structural studies of hTIM-3 also allow for a deeper understanding of these targeting agents. An example of this is the anti-hTIM-3 antibody 2E2 that has immunomodulatory properties and binds to the CC′ and FG loops of the hTIM-3 IgV domain but not hTIM-1 IgV domain (Supplementary Fig. 6a). Consistent with this, peptide competition assays show that a hTIM-3 C-C′ loop-derived peptide, but not a scrambled control peptide, inhibits binding of mouse anti-human 2E2 antibody to hTIM-3 expressed on exhausted CD4+ T-cells or transfected Jurkat T cells (Supplementary Fig. 6b). Our studies provide a potential explanation for this selectivity which may be useful in designing future therapeutic agents to specifically target hTIM-3. Interestingly, our structural comparisons of hTIM-1 and hTIM-3 show that hTIM-1, despite exhibiting a similar conformation of the GFCC′ face, has non-conserved aromatic residues Trp117 and Phe118 at the tip of the hTIM-1 FG loop (Supplementary Fig. 6c). It is predicted that these residues may impose direct steric hindrance on therapeutic antibody binding to hTIM-1, as we predict is the case for the 2E2 antibody.

In conclusion, the high-resolution crystal structure and NMR analysis described herein provide structural mapping of the hTIM-3 IgV domain in the context of its various roles. It is anticipated that these studies will guide future understanding of these interactions and the development of novel therapeutic agents directed at hTIM-3 and those that are currently in human clinical development.

**Methods**

**Protein purification, Crystallization, data collection, structure determination and analysis.** Human TIM-3 IgV domain protein (N-terminal methionine followed by Val24 thru Lys130) was purified using our previously published protocols with slight modifications in refolding conditions. Briefly, refolding of human TIM-3 IgV was performed by drop-wise addition of solubilized inclusion bodies into 1 liter of a refolding buffer consist of 100 mM Tris pH 8.3, 0.4 M L-arginine, 2 mM EDTA, 5 mM reduced GSH and 0.5 mM oxidized GSH and refolding mixture was stirred overnight at 4 °C. Next, the mixture was concentrated to 100 ml and
proteins were expressed and purified as described4. Information to create anomalous difference Fourier (using CCP4 FFT) 28 and LLG (using Phaser-EP “ MR-SAD all hTIM-3 residues, which led to the final crystallographic Rwork
10 cycles of maximum likelihood Restrained refinement using isotropic B factors with “Use automatic weighing ″
Thus, we concluded that Ca++-bound hTIM-3 structure were collected using the LS-CAT beamlines 21-ID-F/G at Advanced Photon Source (APS, Argonne, IL, USA). The data was processed with the HKL200025, *iMosfilm*27 and the CCP4 suite28. During the data collection, the initial space group was determined with HKL2000 software which was available to us at the beaml ine sources (APS, Argonne, IL: NSLS-II, Upton, NY). Once the initial data collection was completed, we determined the space group of the whole data set with *iMosfilm* software which was easily available to us in our laboratory. During the data processing, we set the high-resolution limit at 1.7 Å upon the overall and outer shell completeness. Although the 1/σ(I) value was 10.5 in the highest resolution shell, when we increased the highest resolution shell to 1.6 Å or further, we observed a considerable decrease in completeness of the highest resolution shell. As a result, we embraced a conservative approach and limited the highest resolution shell to 1.7 Å.

The structure of the hTIM-3 IgV domain with Ca++ was determined by a molecular replacement method using a modified N-terminal domain of a another hTIM-3 crystal structure (PDB code 5F71)23 as a search model, with all hTIM-3 amino acids changed to alanine to reduce model bias and all ligands, water and metal ions, were removed. Molecular replacement was performed with Molrep, refinements were performed with the Refmac and intermittent model building was performed with COOT29. 5% of the total reflection data was excluded from the refinement cycles and used to calculate the free Rfactor (Rfree) for monitoring the refinement progress. We used 10 cycles of maximum likelihood Restricted refinement using isotropic B factors with “Use automatic weighing” option turned on. Although some of loops and strands fitted properly during the early refinement process with Rwork/Rfree of 42/45%, significant structural remodeling of entire B-C, C-C,’ C-C′, C-D, and F-G loops, guided by a mTIM-3 crystal structure3 (PDB code 2OYP), were required for better fitting and refinement. An electron density for a metal ion consistent with Ca++ was identified in the crystal structure after rigid body and subsequent refinements associated with model building. Notably, Ca++ was fitted with an occupancy of 1 only after fitting of all hTIM-3 residues, which led to the final crystallographic Rwork/Rfree of 15% at 1.7 Å resolution for a Ca++ bound hTIM-3 structure. An electron density associated with benzoic acid was noted adjacent to the Ca++ binding site that was fit during the refinement. We also tried to fit sodium (Na+) instead of Ca++ for fitting a metal ion density as described above, but an improved structural model, electron density maps and Rfree/Rwork were obtained with Ca++ that included interactions with residues of the hTIM-3 F-G loop similar to that observed in our NMR binding experiments. While Rfactor/Rfree values of 0.1502/0.1805 were observed for Ca++ higher Rfactor/Rfree values of 0.1521/0.1812 were observed for Na+, with a significant positive Fo-Fc map density, supporting Ca++ as the observed metal ion. Given that Ca++ exhibits a significant anomalous signal in contrast to Na+, we used this information to create anomalous difference Fourier (using CCP4 FFT)28 and LLG (using Phaser-EP “ MR-SAD phasing” gui in PHENIX)38 maps. These maps clearly revealed the peak as representative of Ca++ at 5 sigma. Thus, we concluded that Ca++ exhibited a better fit which was consistent with the interactions observed in our NMR studies. The atomic coordinates and structure factors for the Ca++ bound hTIM-3 crystal structure were deposited with RCSB accession code 6DHB.

The X-ray data and structure refinement statistics are shown in Table 1. All the figures were drawn using PyMOL (DeLano Scientific); sequence alignments were done using Clustal Omega39, solvent accessible residues of Ca++-bound hTIM-3 crystal structure reported here were determined by Swiss-PDB viewer using 30% accessibility of residues by solvent; cavity volume (Connolly’s molecular surface) between C-C′ and F-G loop was determined using CASTp webserver31, and mapping of glycosylation sites was done using GPP prediction webserver32.

**Differential scanning fluorimetry.** Differential scanning fluorimetry experiments were performed on a QuantStudio 6 (Applied Biosystems). Samples of hTIM-3 IgV domain (20 µM) were prepared with 0–10 mM CaCl2 and incubated for 4 hours at 4 °C. 200 µL reactions were set up in a 96 well PCR plate comprised of each hTIM-3/CaCl2 mixture, 1000x fluorescent probe (Protein Thermal Shift Dye Kit, ThermoFisher) in triplicate and subjected to heat ramp from 25–99 °C. Fluorescence was measured in 0.25 °C intervals and the melting temperatures (Tm) were calculated from the respective fluorescence profiles using the Boltzmann method in Protein Thermal Shift™ Software (Applied Biosystems). Standard error was calculated from experimental triplicate sample repeats.

**NMR assignments of human TIM3 IgV domain.** 15N and 13C double-labeled hTIM-3 IgV domain protein were expressed from E. coli in M9 minimal medium containing 15NH4Cl and 13C-glucose as the sole nitrogen and carbon sources, and purified as described above. Non-uniformly-sampled (NUS) triple resonance
experiments using $^{15}$N/$^{13}$C-hTIM-3 IgV (0.85 mM) in 10 mM HEPES, 50 mM NaCl, pH 7.0 with 10% D$_2$O, were performed at 25 °C on a 700 MHz Agilent DD2 spectrometer equipped with a cryogenic probe. The data were processed using NMRPipe$^{33}$ and Iterative Soft Thresholding reconstruction approach (istHMS)$^{34}$ and analyzed by CARA$^{35}$. Overall 75% of backbone amide resonances were assigned.$^{34}$ The missing assignments in the $\alpha$, $\beta$ and $\gamma$ strand regions are likely caused by conformation exchange dynamics in solution in the absence of human TIM3 ligands. Secondary structure predictions based on assigned chemical shifts were obtained using the TALOS+ software$^{36}$. The NMR data and assigned chemical shifts were deposited in the BMRB database with accession code 27525.

**NMR titration study of human TIM3 IgV domain.** Calcium titration experiments were performed using 0.05 mM $^{15}$N-hTIM-3 IgV domain purified protein titrated with 0 mM to 100 mM CaCl$_2$ on a 600 MHz Agilent DD2 spectrometer at 25 °C, processed by NMRPipe and analyzed by CARA. An initial estimate of the dissociation constant of Ca$^{++}$ binding to the un-ligated hTIM-3 IgV domain yielded $K_d$ of 22 mM by fitting an approximate linear solution to the reciprocal plot of the Phe61 chemical shift relative to the Ca$^{++}$ concentration. More rigorous non-linear regression analysis of the dissociation constants were further performed using Microsoft Excel by fitting the chemical shift change index $\delta = \sqrt{(\Delta Ncs/5^2 + (\Delta Hcs)^2)}$ relative to the Ca$^{++}$ concentration according to the equation provided by Morton et al.$^{37}$ By non-linear regression analysis, the average dissociation constant of Ca$^{++}$ binding to the un-ligated hTIM-3 IgV was determined as 27.8 mM for five amide peaks with the largest observable shift changes being 26.6 mM, 27.2 mM, 24.6 mM, 29.5 mM and 31.0 mM for Leu48, Phe61, Arg89, Arg111 and Gln113 respectively.

**ELISA binding and blockade experiments.** ELISA binding studies were performed using Costar 96-well high protein binding plates (Corning). For binding studies between human TIM-3 Fc fusion protein (BPS Bioscience) and tagless hCEACAM1 IgV domain, 50 μl of 250 nM tagless hCEACAM1 protein in Tris-buffered saline buffer containing 10 mM CaCl$_2$ (TBS-Ca$^{++}$) or TBS-Ca$^{++}$ buffer alone was added to wells of an ELISA plate and incubated overnight at 4 °C. After incubation overnight, the wells were washed three times with TBS buffer containing 0.05% Tween 20 and 10 mM CaCl$_2$ (TBST-Ca$^{++}$), blocked with 2% BSA in TBST-Ca$^{++}$ buffer for 1 hour at room temperature, and incubated with hTIM-3 Fc fusion protein (0–4 μM) or hIgG fragment protein (Abcam) in TBST-Ca$^{++}$ buffer for two hours at 37 °C. The wells were washed three times with TBST-Ca$^{++}$ buffer and incubated with goat anti-human IgG-HRP conjugated antibody (Southern Biotech) for 1 hour at room temperature. The wells were washed three times with TBST-Ca$^{++}$ buffer and treated with TMB peroxidase substrate solution (Seracare). The reaction was stopped using 3 M HCl and OD values were obtained at 450 nm. Data were plotted using GraphPad Prism software and EC$_{50}$ was determined. Similar ELISA protocols were used for peptide blocking experiments where binding of tagless hCEACAM1 was measured with hTIM-3 Fc fusion protein (4 μM) incubated with 10 μM of hTIM-3 C-C′′′′ loop-derived peptide (amino acids 58–77, 58-CPVFECCNVVLTIDERDVY(77)) or 10 μM of scrambled control peptide (TLTCCFVNPYDVRVNDRE) in Tris-buffered saline buffer containing 10 mM CaCl$_2$. The peptides were synthesized by American Peptide Company, CA and with >98% purity. Data were plotted using GraphPad Prism software and $p$ values determined using 1 way Anova-Tukey's multiple comparisons test, where the mean of each binding column was compared with the mean of every binding column. For ELISA binding experiments between tagless human TIM-3 protein used for structural studies and GST tagged hCEACAM1 IgV-domain protein, 50 μl of 250 nM tagless hTIM-3 protein was added to wells of ELISA plates using similar ELISA protocols described above in Dulbeccco PBS buffer (Sigma Aldrich) and binding of GST tagged hCEACAM1 IgV domain protein (0–15 μM), GST control protein (Sigma Aldrich), was detected using goat polyclonal anti-GST-HRP conjugated antibody (Abcam). Binding data were plotted and EC$_{50}$ was determined in Graphpad Prism. Blockade of tagless hTIM-3 binding to GST tagged CEACAM1 IgV domain protein was performed in the presence of a mouse anti-human CEACAM1 IgV domain specific monoclonal, SF4 (0–1 μM), or mouse IgG1 isotype control antibody, using 4 μM GST-tagged hCEACAM1 IgV domain protein.

**Peptide competition assays.** For in vitro CD4$^+$ T cell activation to induce exhaustion, MACS purified human CD4$^+$ T cells were cultured in complete medium RPMI 1640 (Lonza) supplemented with 10% fetal calf serum (FCS), 1% glutamine, 100 IU/ml penicillin, 100 μg/ml streptomycin (Life Technologies), 25 mM HEPES (Sigma-Aldrich) in 96-well plates at a concentration of 40 IU/ml recombinant (r)IL-2 (NIH) and soluble anti-CD3 (1 μg/ml) in a total volume of 200 ml. Cell cultures were refreshed with new rIL-2 and new soluble anti-CD3 every 5 days. After three rounds of stimulation, exhausted CD4$^+$ T cells were generated. Jurkat hTIM3 transfectants have been previously described$^{37}$. The samples from the exhausted CD4$^+$ T cells and Jurkat hTIM3 transfectants were solubilized by suspension in Laemmli sample buffer and the proteins resolved by SDS–PAGE under reducing conditions followed by transfer to a PVDF (polyvinylidene difluoride) membrane. After blocking with 5% of skim milk in 0.05% PBS-Tween (PBS-T) buffer, the membranes were incubated for 16 hr at 4 °C with mouse anti-human 2E2 antibody in the presence of 15-fold molar excess of a hTIM-3 C-C′′′′ loop-derived peptide or scrambled control peptide. The membranes were incubated with goat anti-mouse IgG HRP conjugated antibody (Southern Biotech) for 1 hr at room temperature and visualized by Amersham ECL Western Blotting Detection Reagents (GE Healthcare).

**References**

1. Zhu, C. et al. The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nat Immunol* **6**, 1245–1252 (2005).
2. Dekruyff, R. H. et al. T cell/transmembrane, Ig, and mucin-3 allelic variants differentially recognize phosphatidylserine and mediate phagocytosis of apoptotic cells. *J Immunol* **184**, 1918–1930 (2010).
3. Sakuishi, K., Jayaraman, P., Behar, S. M., Anderson, A. C. & Kuchroo, V. K. Emerging Tim-3 functions in antimicrobial and tumor immunity. *Trends Immunol* **32**, 345–349 (2011).
4. Huang, Y. H. et al. CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. Nature 517, 386–390 (2015).
5. Sabatos-Peyton, C. A. et al. Blockade of Tim-3 binding to phosphatidylserine and CEACAM1 is a shared feature of anti-Tim-3 antibodies that have functional efficacy. Oncoimmunology 7, e1385690 (2018).
6. Anderson, A. C. Tim-3: An Emerging Target in the Cancer Immunotherapy Landscape. Cancer Immunol. Res. 2, 393–398 (2014).
7. Jin, H. T. et al. Cooperation of Tim-3 and PD-1 in CD8 T-cell exhaustion during chronic viral infection. Proc. Natl Acad. Sci. USA 107, 14733–14738 (2010).
8. Monney, L. et al. Th1-specific cell surface protein Tim-3 regulates macrophage activation and severity of an autoimmune disease. Nature 415, 536–541 (2002).
9. Cao, E. et al. T cell immunoglobulin mucin-3 crystal structure reveals a galectin-9–independent ligand-binding surface. Immunity 26, 311–321 (2007).
10. Yuan, S. et al. TIM-1 acts as a dual-attachment receptor for Ebolavirus by interacting directly with viral GP and the PS on the viral envelope. Protein Cell 6, 814–824 (2015).
11. Santiago, C. et al. Structures of T cell immunoglobulin mucin protein 4 show a metal-ion–dependent ligand binding site where phosphatidylserine binds. Immunity 27, 941–951 (2007).
12. Sabatos, C. A. et al. Interaction of Tim-3 and Tim-3 ligand regulates T helper type 1 responses and induction of peripheral tolerance. Nat. Immunol. 4, 1102–1110 (2003).
13. Hastings, W. D. et al. TIM-3 is expressed on activated human CD4+ T cells and regulates Th1 and Th17 cytokines. Eur. J. Immunol. 39, 2492–2501 (2009).
14. Sakushi, K. et al. Targeting Tim-3 and PD-1 pathways to reverse T cell exhaustion and restore anti-tumor immunity. J. Exp. Med. 207, 2187–2194 (2010).
15. Gao, X. et al. TIM-3 expression characterizes regulatory T cells in tumor tissues and is associated with lung cancer progression. PLoS One 7, e30676 (2012).
16. Banerjee, H. & Kane, L. P. Immune regulation by Tim-3. F1000Res. 14 (2018).
17. Lee, J. et al. Phosphotyrosine-Dependent Coupling of Tim-3 to T-Cell Receptor Signaling Pathways. Mol. Cell. Biol. 31, 3963–3974 (2011).
18. Van de Weyer, P. S. et al. A highly conserved tyrosine of Tim-3 is phosphorylated upon stimulation by its ligand galectin-9. Biochem. Biophys. Res. Commun. 351, 571–576 (2006).
19. Yanagihashi, Y., Segawa, K., Maeda, R., Nabeshima, Y. & Nagata, S. Mouse macrophages show different requirements for phosphatidylserine receptor Tim4 in effecrocytosis. Proc. Natl. Acad. Sci. 114, 8800–8805 (2017).
20. Chiba, S. et al. Tumor-infiltrating DCs suppress nucleic acid-mediated innate immune responses through interactions between the receptor Tim-3 and the alarmin HMGB1. Nat. Immunol. 13, 832–842 (2012).
21. Santiago, C. et al. Structures of T Cell immunoglobulin mucin receptors 1 and 2 reveal mechanisms for regulation of immune responses by the TIM receptor family. Immunity 26, 299–310 (2007).
22. Sharma, D. & Rajarathnam, K. 13C NMR chemical shifts can predict disulfide bond formation. J. Biomol. NMR 18, 165–171 (2000).
23. Wang, H. et al. Crystal structures of human TIM members: Ebolavirus entry-enhancing receptors. Chin. Sci. Bull. 60, 3438–3453 (2015).
24. Morales, V. M. et al. Regulation of human intestinal intraepithelial lymphocyte cytolytic function by biliary glycoprotein (CD66a). J. Immunol. 163, 1363–1370 (1999).
25. McIntyre, J. et al. Identification of Tapr (an airway hyperreactivity regulatory locus) and the linked Tim gene family. Nat. Immunol. 2, 1109–1116 (2001).
26. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. Methods in Enzymology 307, 307–326 (1997).
27. Batty, T. E. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 271–281 (2011).
28. Winn, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242 (2011).
29. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501 (2010).
30. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 11, 539 (2015).
31. Binkowski, T. A., Naghibzadeh, S. & Liang, J. CASTp: Computed Atlas of Surface Topography of proteins. Nucleic Acids Res. 31, 3352–3355 (2003).
32. Hanby, S. E. & Hirst, J. D. Prediction of glycosylation sites using random forests. BMC Bioinformatics 9, 500 (2008).
33. Delaglio, F. et al. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293 (1995).
34. Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H. & Wagner, G. Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. J. Biomol. NMR 52, 315–327 (2012).
35. Keller, R. Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment. Swiss Inst. Fed. Technol. Zurich (2004).
36. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biomol. NMR 44, 213–223 (2009).
37. Morton, C. J. et al. Solution structure and peptide binding of the SH3 domain from human Fyn. Structure 15(4(6)), 705–14 (1996).
38. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Cryst. D66, 213–221 (2010).

Acknowledgements

We thank the staff of NSLS-II (Upton, NY, USA) and LS-CAT beamlines 21-ID-F and 21-ID-G (APS; Argonne, Illinois, USA) for x-ray data collection and A. Russell for assistance in purification of proteins used in NMR study. We are grateful to M. Pyzik, J. Engen, and T. Wales for helpful discussions. This work was supported by the NIH Grants 5R01DK051362-21 and the High Pointe Foundation to R.S.B., 5P01AI073748-09 to V.K.K., and GM047467 and AI037581 to G.W.

Author Contributions

A.K.G., W.M.K. and Z.-Y.J.S. performed most experiments. A.K.G. performed crystallization and structure determination, Z.-Y.J.S. carried out NMR experiments and W.M.K. conducted biophysical assessments. A.K.G., Y.-H.H., and Y.K. designed ELISA and peptide competition experiments. D.A.B. and E.J.S. designed plasmid DNA constructs and established protein refolding protocol. G.P. provided expertise in X-ray crystallography and structural comparison. G.W. provided expertise in NMR studies. R.S.B. and V.K.K. established this project, and together with A.K.G., W.M.K., Z.-Y.J.S., Y.-H.H. and Y.K. wrote the manuscript. R.S.B. is the senior author of this paper.
Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-35754-0.

Competing Interests: R.S.B. and V.K.K. are consultants to Syntalogic Pharmaceuticals and R.S.B. has an equity interest in Syntalogic Pharmaceuticals, which is developing therapeutics to target CEACAM1.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018