Structure of the N-terminal dimerization domain of CEACAM7
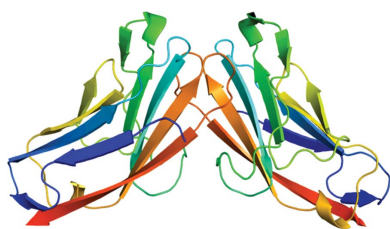
Daniel A. Bonsor,[a] Dorothy Beckett[b] and Eric J. Sundberg[a,c,d]*

[a]Institute of Human Virology, University of Maryland School of Medicine, Baltimore, MD 21201, USA, [b]Department of Chemistry and Biochemistry, University of Maryland College Park, Baltimore, MD 20742, USA, [c]Department of Medicine, University of Maryland School of Medicine, Baltimore, MD 21201, USA, and [d]Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201, USA. *Correspondence e-mail: esundberg@som.umaryland.edu

CEACAM7 is a human cellular adhesion protein that is expressed on the surface of colon and rectum epithelial cells and is downregulated in colorectal cancers. It achieves cell adhesion through dimerization of the N-terminal IgV domain. The crystal structure of the N-terminal dimerization domain of CEACAM has been determined at 1.47 Å resolution. The overall fold of CEACAM7 is similar to those of CEACAM1 and CEACAM5; however, there are differences, the most notable of which is an insertion that causes the C'' strand to buckle, leading to the creation of a hydrogen bond in the dimerization interface. The $K_{dimerization}$ for CEACAM7 determined by sedimentation equilibrium is tenfold tighter than that measured for CEACAM5. These findings suggest that the dimerization affinities of CEACAMs are modulated *via* sequence variation in the dimerization surface.

1. Introduction

Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) belong to the immunoglobulin (Ig) family and are expressed differentially on the surfaces of cells (Gray-Owen & Blumberg, 2006; Tchoupa *et al.*, 2014). There are 12 CEACAMs found in humans: CEACAM1, CEACAM3–CEACAM8, CEACAM16 and CEACAM18–CEACAM21 (Beauchemin & Arabzadeh, 2013; Tchoupa *et al.*, 2014). Their functions and roles in cellular processes are diverse and include roles in phagocytosis, hearing, proliferation, signaling, tumor suppression and cell adhesion (Oikawa *et al.*, 1989, 1991; Benchimol *et al.*, 1989; Streichert *et al.*, 2001; Pils *et al.*, 2008; Singer *et al.*, 2010; Zheng *et al.*, 2011). CEACAMs are typically dysregulated in cancer and are found to be parasitized by bacteria (*e.g. Neisseria meningitidis*, *Escherichia coli* and *Haemophilus influenzae*) and viruses in mice (*e.g.* coronavirus) during infection (Dveksler *et al.*, 1991; Leusch *et al.*, 1991; Bos *et al.*, 1997; Schölzel *et al.*, 2000; Virji *et al.*, 2000; Duxbury *et al.*, 2004; Litkouhi *et al.*, 2008; Obrink, 2008; Singer *et al.*, 2010). CEACAMs contain an N-terminal immunoglobulin variable domain (IgV), a variable number of immunoglobulin constant domains (IgC2) and either a C-terminal transmembrane and cytoplasmic domain or a glycophosphatidyl-inositol (GPI) moiety by which they are anchored to the plasma membrane (Tchoupa *et al.*, 2014). Cell adhesion is achieved through the N-terminal domain of CEACAMs, which can undergo heterodimerization and homodimerization in a *cis* (on the same cell) or *trans* (across different cells) fashion (Taheri *et al.*, 2000; Kuroki *et al.*, 2001; Watt *et al.*, 2001).

Table 1
CEACAM7 production details.

Source organism	*Homo sapiens*
DNA source	Synthetic
Expression vector	pET-21d
Expression host	*E. coli* BL21 (DE3) pLysS
Complete amino-acid sequence of the construct produced	MAQTNIDVVPFNVAEGKEVLLVVHNESQNLYGYN-WYKGERVHANYRIIGYVKNISQENAPGPAHNG-RETIYPNGTLLIQNVTHNDAGIYTLHVIKENL-VNEEVTRQFYVF

Table 2
CEACAM 7 crystallization.

Method	Hanging drop
Plate type	VDX48
Temperature (K)	293
Protein concentration (mg ml^{-1})	5.9
Buffer composition of protein solution	20 mM Tris–HCl, 100 mM NaCl pH 7.5
Composition of reservoir solution	18% PEG 8000, 0.1 M Tris–HCl pH 8.5, 0.2 M magnesium chloride
Volume and ratio of drop	2 µl, 1:1
Volume of reservoir (µl)	200

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline 23-ID-B, APS
Wavelength (Å)	1.0332
Temperature (K)	100
Detector	MAR Mosaic 300 mm CCD
Crystal-to-detector distance (mm)	253.8
Rotation range per image (°)	0.2
Total rotation range (°)	180
Exposure time per image (s)	0.3
Space group	$P2_1$
a, b, c (Å)	32.62, 64.89, 103.16
α, β, γ (°)	90, 89.99, 90
Mosaicity (°)	0.52
Resolution range (Å)	103.16–1.47 (1.50–1.47)
Total No. of reflections	205245
No. of unique reflections	71599
Completeness (%)	98.2 (82.7)
Multiplicity	2.900 (1.80)
$\langle I/\sigma(I)\rangle$	28.0 (1.57)†
$R_{\mathrm{r.i.m.}}$‡	0.130 (0.666)
Overall B factor from Wilson plot (Å^2)	18.3
CC$_{1/2}$	0.993 (0.664)

† The data were extended owing to a reasonable CC$_{1/2}$ value. An $\langle I/\sigma(I)\rangle$ of 2.0 is equal to a resolution of 1.57 Å. ‡ Estimated $R_{\mathrm{r.i.m.}} = R_{\mathrm{merge}}[N/(N-1)]^{1/2}$, where N is the data multiplicity.

CEACAM7 is expressed on highly differentiated epithelial cells of the colon and rectum and on the epithelial cells within the ducts of the pancreas (Schölzel *et al.*, 2000). The expression pattern of CEACAM7 suggests a specialized function. In fetal tissues of the colon, CEACAM7 is located at the base of epithelial cells and has been found to migrate to the apical surface a few days after birth (Schölzel *et al.*, 2000). CEACAM7 contains three domains: an N-terminal IgV domain, a single IgC2 domain and a cell-surface GPI anchor domain (Tchoupa *et al.*, 2014). CEACAM7 expression is downregulated during the early development of colorectal tumors, unlike CEACAM5 or CEACAM6, which are typically upregulated, suggesting a tumor-suppression function (Schölzel *et al.*, 2000). Currently, it is unknown whether CEACAM7 is involved in cell adhesion through homo-dimerization or if any structural differences exist that could potentially allow CEACAM7 to function as a tumor-suppression molecule when compared with the two known structures of CEACAM1 and CEACAM5 (Fedarovich *et al.*, 2006; Korotkova *et al.*, 2008). Here, we report the 1.47 Å resolution X-ray crystal structure of the N-terminal domain of CEACAM7 and have characterized its oligomeric state in solution.

2. Materials and methods

2.1. CEACAM7 production

The N-terminal domain of CEACAM7 was synthesized as a codon-optimized GeneArt string (Life Technologies), which was digested and ligated into an NcoI/XhoI-cut pET-21d vector without a purification tag. CEACAM7 was expressed in inclusion bodies in *E. coli* BL21 (DE3) pLysS cells. Briefly, 1 l of cells were grown in LB Miller at 310 K until an OD$_{600\,\mathrm{nm}}$ of ~0.6 was attained, prior to induction with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were grown for a further 4 h before harvesting (5000g for 15 min at 277 K). The

cells were resuspended in lysis buffer [50 mM Tris–HCl, 500 mM NaCl, 1%(v/v) Triton X-100 pH 7.5] and lysed by sonication. Inclusion bodies were isolated (20 000g for 20 min at 277 K), resuspended in lysis buffer, sonicated and isolated by centrifugation (20 000g for 20 min at 277 K). Inclusion bodies were washed with a high-salt buffer (50 mM Tris–HCl, 1.0 M NaCl pH 8.0) to remove DNA, followed by lysis buffer without Triton X-100. Inclusion bodies were dissolved in 30 mM Tris–HCl, 150 mM NaCl, 8.0 M urea pH 8.3 (~5 ml per litre of grown cells), refolded by rapid dilution (1:12 ratio) at 277 K into 50 mM CHES–HCl, 500 mM L-arginine pH 9.2 and left for 24 h. Refolded CEACAM7 was dialyzed against 10 mM Tris–HCl pH 8.0 and concentrated by anion-exchange chromatography (Mono Q, GE Healthcare). A linear salt gradient from 0 to 1000 mM was run at 1 ml min^{-1} over 15 min, with CEACAM7 eluting at between 50 and 100 mM NaCl. CEACAM7 was further purified by size-exclusion chromatography (Superdex 200, GE Healthcare) in 50 mM Tris–HCl, 150 mM NaCl, 1 mM EDTA pH 7.5 and fractions were stored at 277 K. Typically, 10 mg of refolded protein per litre was obtained with a refolding efficiency of 10%. Macromolecule-production information is summarized in Table 1.

2.2. CEACAM7 crystallization

CEACAM7 was concentrated using a Centricon centrifugal filter unit (10 kDa MWCO, Millipore) and subsequently dialyzed against 20 mM Tris–HCl, 100 mM NaCl pH 7.5. CEACAM7 at 5.9 mg ml^{-1} was screened against The JCSG+ Suite screen (Qiagen) using a Crystal Gryphon Protein Crystallography System (Art Robbins Instruments) with sitting drops consisting of 150 nl protein solution and 150 nl reservoir solution equilibrated against 50 µl reservoir solution. A shower of small crystals grew over 5 d in condition D6

Table 4
CEACAM7 structure refinement.

Resolution range (Å)	103.16–1.47 (1.50–1.47)
Completeness (%)	97.9
σ Cutoff	$F > 0.000\sigma(F)$
No. of reflections, working set	67967 (4156)
No. of reflections, test set	3613 (226)
Final R_{cryst}	0.143 (0.280)
Final R_{free}	0.193 (0.309)
Cruickshank DPI	0.0722
No. of non-H atoms	
Protein	3529
Ion	3
Ligand	0
Water	269
Total	3801
R.m.s. deviations	
Bonds (Å)	0.020
Angles (°)	1.901
Average B factors (Å^2)	
Protein	27.6
Ion	41.9
Ligand	0.0
Water	39.5
Ramachandran plot	
Favored regions (%)	98.8
Additionally allowed (%)	0.9
PDB code	4y89

[20%(w/v) polyethylene glycol 8000, 200 mM magnesium chloride, 100 mM Tris–HCl pH 8.5]. Crystals were optimized by hanging-drop vapor diffusion, with the final crystals growing in 18%(w/v) polyethylene glycol 8000, 200 mM magnesium chloride, 100 mM Tris–HCl pH 8.5. Crystallization information is summarized in Table 2.

2.3. Data collection, processing, structure solution and refinement

Crystals of CEACAM7 were washed and cryoprotected in mother liquor containing 20%(v/v) glycerol. Data were collected on beamline 23-ID-B at the Advanced Photon Source (APS), Argonne National Laboratory, USA. Data were processed using *HKL*-2000 (Otwinowski & Minor, 1997). Data-collection and processing statistics are shown in Table 3. F_{obs} were obtained using *SCALEPACK*2*MTZ* (Winn *et al.*, 2011). Molecular replacement was performed using *MOLREP* (Vagin & Teplyakov, 2010) and a *CHAINSAW* (Stein, 2008) model of CEACAM5 (PDB entry 2qsq; Korotkova *et al.*, 2008), a protein with 65% sequence identity to CEACAM7. CEACAM7 was refined with *REFMAC* (Murshudov *et al.*, 2011) and rebuilt in *Coot* (Emsley *et al.*, 2010). *MolProbity* (Chen *et al.*, 2010) was used for Ramachandran analysis. Refinement statistics are shown in Table 4.

2.4. Analytical ultracentrifugation

Sedimentation-equilibrium measurements of CEACAM7 were performed using a Beckman–Coulter XL-I analytical ultracentrifuge equipped with a four-hole An-60 Ti rotor at 20°C. Prior to centrifugation, CEACAM7 was dialyzed extensively against 50 mM Tris–HCl pH 7.5, 50 mM NaCl. *SEDNTERP* (http://sednterp.unh.edu) was used to calculate values for the protein partial specific volume and solvent

density from the protein amino-acid sequence and buffer composition, respectively. CEACAM7 at three different concentrations (24.1, 14.5 and 9.6 µM) was loaded into cells equipped with six-hole charcoal-filled Epon centerpieces (1.2 cm path length) with sapphire windows. Centrifugation was carried out at 29 000, 32 000 and 35 000 rev min^{-1} and scans were acquired at 280 nm with a step size of 0.001 and five averages per step. The data were globally analyzed using the *WinNonLin* program (Johnson *et al.*, 1981).

3. Results and discussion

A single X-ray diffraction data set was collected to a resolution of 1.47 Å. Initial indexing of the data suggested that the crystal contained a primitive orthorhombic lattice with two molecules in the asymmetric unit. However, attempts to find a molecular-replacement solution using the CEACAM5 monomer as a search model yielded no solutions that would refine in any of the orthorhombic space groups. The data were reprocessed in a primitive monoclinic lattice with a β angle of 89.99°, which led to the correct solution in space group $P2_1$ with four copies of the search model found in the asymmetric unit. $\langle |L| \rangle$ tests of the data (0.503) show that the data are untwinned and no pseudo-merohedral twinning was detected. All residues in the final model were modeled except for the initial alanines of two of the four molecules in the asymmetric unit. The final model contained three chloride ions and 269 waters. The final model was refined to an R_{cryst} and R_{free} of 0.143 and 0.193, respectively. The structure factors and model have been deposited in the Protein Data Bank (PDB entry 4y89).

The closest homologs of CEACAM7 in the PDB are CEACAM1 and CEACAM5, which both share 65% sequence identity with CEACAM7, with 38 residues differing between the proteins (Fig. 1a). The overall fold of CEACAM7 is similar to those of the other CEACAMs that have been determined previously. The overall topology is that of the V-set fold of the immunoglobulin superfamily, comprised of two β-sheets labeled *ABED* and *A'GFCC'C''*. The sheets are connected by the *BC*, *EF*, *C''D* and *AA'* loops (Fig. 1b). The r.m.s.d.s of the CEACAM7 molecules to each other are low (~0.30 Å), showing no structural differences within the asymmetric unit. The four molecules of CEACAM7 form two pairs of dimers (Fig. 1c). The dimer interface is formed from the second β-sheet, *A'GFCC'C''*, specifically the *GFCC'C''* strands and the *CC'*, *C'C''* and *FG* loops (Fig. 1c). Dimerization buries 1610 Å^2 of solvent-accessible surface area as calculated by *PISA* (Krissinel & Henrick, 2007). This is similar to the CEACAM1 and CEACAM5 homodimers, which bury 1600 and 1460 Å^2 of solvent-accessible surface area, respectively. The shape-complementarity value (Sc; Lawrence & Colman, 1993) is 0.68, which is smaller than those for the other CEACAMs, with values of 0.81 and 0.72 for CEACAM1 and CEACAM5, respectively. CEACAM7 forms nine hydrogen bonds in the dimerization interface. This is more than CEACAM5 (six hydrogen bonds) but less than CEACAM1 (16 hydrogen bonds). Of the 38 residues in CEACAM7 that

differ from CEACAM1 and CEACAM5, eight are found in the dimerization interface.

The dimerization constant of CEACAM5 was determined previously to be $0.8\,\mu M$ by analytical ultracentrifugation (Korotkova *et al.*, 2008). CEACAM1 does dimerize but forms high molecular-weight oligomers (Korotkova *et al.*, 2008). The dimerization constant of CEACAM7 was estimated by sedimentation-equilibrium analysis using analytical ultra-
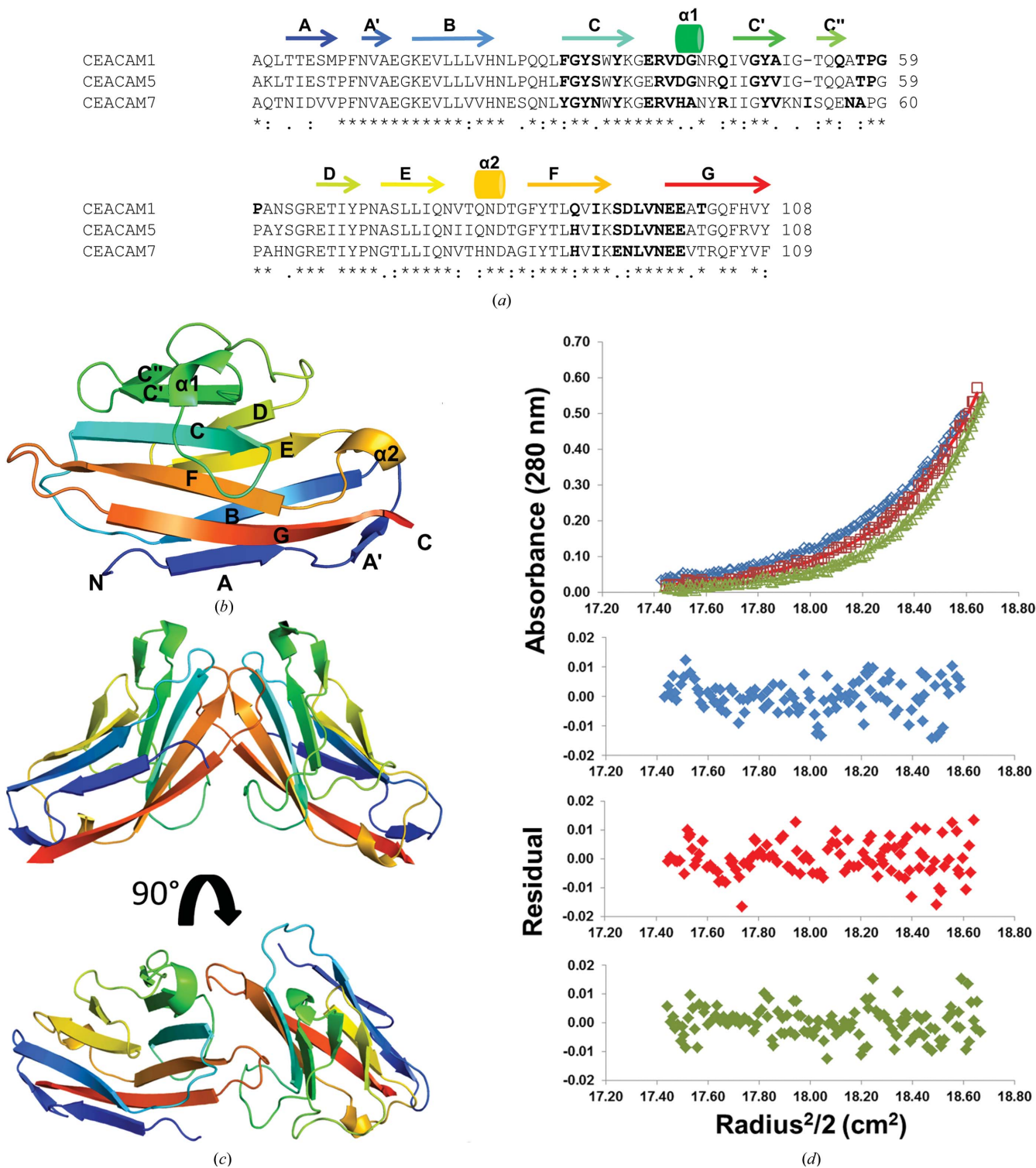

Figure 1
(*a*) Sequence alignment of CEACAM1, CEACAM5 and CEACAM7. Residues buried in the dimerization interface are shown in bold. (*b*) Overall topology of the CEACAM7 fold. (*c*) Side and top view of the CEACAM7 dimer. (*d*) Top, sedimentation-equilibrium experiment of CEACAM7 (28 μ*M*) in 50 m*M* Tris–HCl, 50 m*M* sodium chloride pH 7.5 with rotor speeds of 29 000, 32 000 and 35 000 rev min^{-1} (blue, red and green curves, respectively). Bottom, residuals of fitted data for each curve.

centrifugation (Fig. 1*d*). An estimated average molecular weight of 24.7 ± 0.7 kDa (the theoretical monomer molecular weight is 12 574 Da) and a $K_{dimerization}$ of 95 n*M* (+20/−60 n*M*) were measured, a tenfold increase in affinity when compared with CEACAM5. We observed no higher molecular weight oligomers for CEACAM7 other than the dimer.

Superposition of the CEACAM7 dimer (*A* + *B*) onto the CEACAM5 dimer (*A* + *B*) was achieved using one half of each dimer (molecule *A*) and two r.m.s.d.s were calculated: one for the first half of the dimer (molecule *A*), which results in an r.m.s.d. of 0.67 Å, and the second for the second half, which shows a larger r.m.s.d. of 2.70 Å (molecule *B*; Fig. 2*a*). Superposition of CEACAM7 onto the CEACAM1 dimer using the same method reveals similar r.m.s.d.s of 0.83 Å for molecule *A* and 2.26 Å for molecule *B* (Fig. 2*b*). Closer comparison of CEACAM7 with CEACAM5 and CEACAM1 shows two major regions of deviation. The first is the *BC* loop (residues 23–29), which is not involved in the dimerization interface (Fig. 2*c*). This region is highly conserved among members of the CEACAM family; however, positions 25–26 of CEACAM7 differ from those of the other CEACAMs. In all other CEACAMs these residues are Leu25 and Pro26. However, in CEACAM7 they are Glu25 and Ser26. The loss of Pro26 is likely to reduce the rigidity of the loop. The r.m.s.d. of this loop in CEACAM7 is 3.10 Å relative to CEACAM5 and

the displacement of this loop causes a slight movement of the N-terminal β-strand. Notably, the preceding residue is Asn24, a unique N-linked glycosylation site found only in CEACAM7 and CEACAM4. Three other glycosylation sites are present in CEACAM7 (Asn52, Asn72 and Asn79). These are highlighted in Fig. 2(*d*), showing that none are found in the dimerization interface. Glycosylation of CEACAM5 has been shown to be important for interaction with CD8α (Roda *et al.*, 2014) and therefore may also be important for CEACAM7 function.

The second region of deviation between CEACAM7, CEACAM5 and CEACAM1 is the *C'C''* loop and the *C''* strand. CEACAM7 is unique compared with other CEACAMs, as the *C'C''* loop contains a single amino-acid insertion (isoleucine) between residues 52 and 53. To accommodate the insertion in the *C'C''* loop without altering the length of the *C'C''* or the *C''D* loops, the *C''* strand is found to be distorted relative to CEACAM5 (Fig. 3*a*) and CEACAM1 (Fig. 3*b*). The *C''* strand bulges in the center (residues 56–60), causing breakages of hydrogen bonds in the antiparallel β-sheet between the *C'* and *C''* strands. The *C''* strand of CEACAM7 is held in place through a hydrogen bond from the $O^{\varepsilon 2}$ atom of Asn57 to the N atom of Gly48 (Fig. 3*c*). In CEACAM5 the carbonyl group of Thr57 forms the hydrogen bond to Gly48 (Fig. 3*d*). Although the *C''* strand does not buckle in CEACAM1, as in CEACAM5, it is found that Thr57
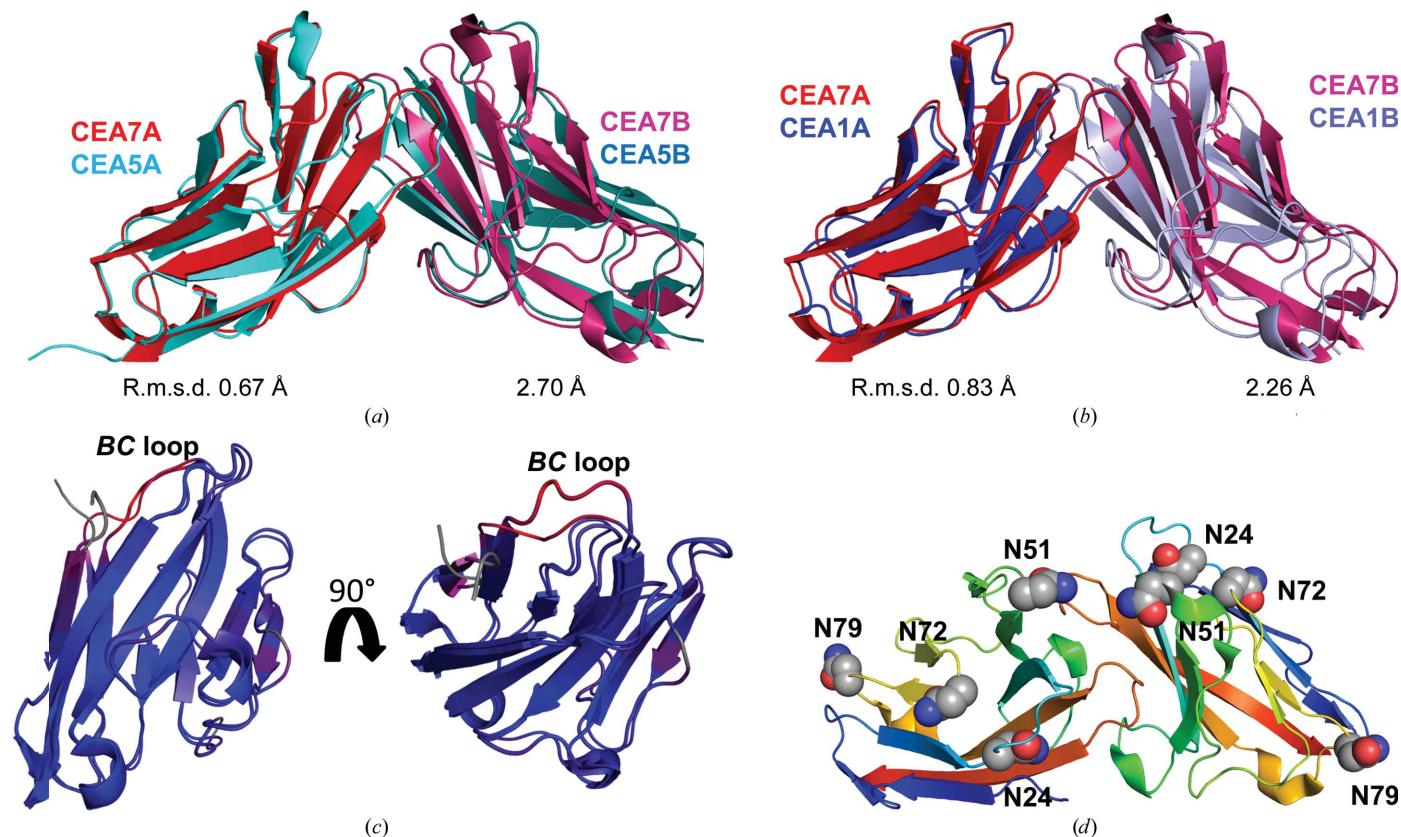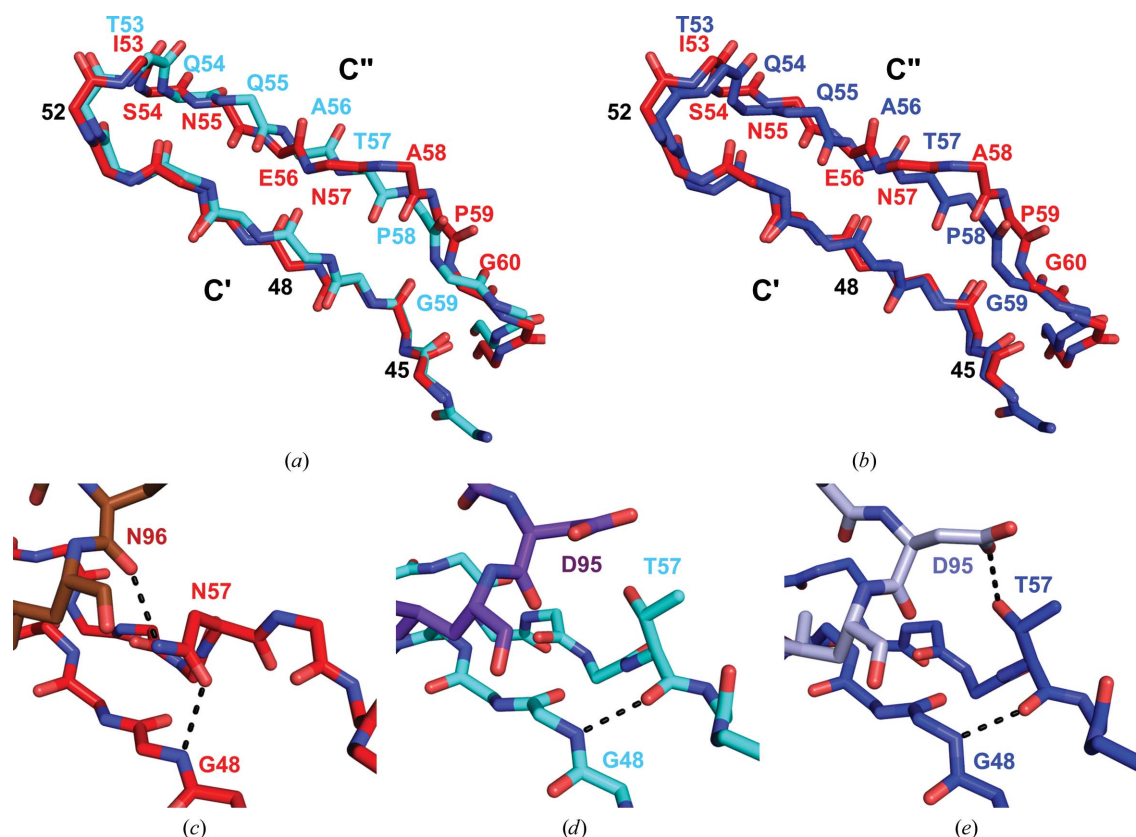

Figure 2
(*a*) Superposition of the CEACAM7 dimer (CEA7A and CEA7B; red and pink, respectively) onto the CEACAM5 dimer (CEA5A and CEA5B; light and dark cyan, respectively) through molecule *A*. (*b*) Superposition of the CEACAM7 dimer (CEA7A and CEA7B; red and pink, respectively) onto the CEACAM1 dimer (CEA1A and CEA1B; blue and lilac, respectively) through molecule *A*. (*c*) Alignment of CEACAM7 and CEACAM5 monomers, colored by r.m.s.d. Dark blue is low r.m.s.d. and red is high r.m.s.d. CEACAM1 is omitted for clarity. (*d*) Potential N-linked glycosylation sites of the CEACAM7 dimer. All residues are solvent-exposed and are not found in the dimerization interface.

Figure 3
(*a*) The polypeptide backbone of the *C′* and *C″* strands of CEACAM7 (red) and CEACAM5 (cyan). (*b*) The polypeptide backbone of the *C′* and *C″* strands of CEACAM7 (red) and CEACAM1 (blue). (*c*) The insertion of Ser54 in CEACAM7 (red) results in breakage of the main-chain antiparallel hydrogen bonds and replacement with a hydrogen bond from the side chain of Asn57. This residue also forms a hydrogen bond across the dimer interface to Asn96 (brown). (*d*) In CEACAM5 (green), only the main-chain antiparallel hydrogen bonds exist. Thr57 is too short to form a hydrogen bond across the dimer interface to Asp95 (purple). (*e*) However, this is not the case for CEACAM1 (blue). Thr57 forms a hydrogen bond across the dimer interface to Asp95 (gray).

does form a hydrogen bond across the dimerization interface to Asp95 (Fig. 3*e*). This buckling of the *C″* strand creates an extra interaction in the dimerization interface, potentially explaining why CEACAM7 forms such tight dimers and has not been found to form heterodimeric CEACAM complexes such as CEACAM6–CEACAM8, CEACAM1–CEACAM8, CEACAM1–CEACAM6, CEACAM3–CEACAM6 and CEACAM5–CEACAM6 (Oikawa *et al.*, 1991; Kuroki *et al.*, 2001; Skubitz & Skubitz, 2008; Singer *et al.*, 2014). Although CEACAM1 does not buckle, it too creates an extra interaction across the dimerization through reorientation of Asp95. This also suggests that this hydrogen bond is important for a higher affinity interaction.

The structure of CEACAM7 reveals that the dimerization interface is comprised of the same face as other CEACAMs (*GFCC′C″*) yet can accommodate 16 different residues (eight from each monomer), suggesting that these sequence differences can modulate the homodimerization to achieve a tenfold increase in affinity.

References

Beauchemin, N. & Arabzadeh, A. (2013). *Cancer Metastasis Rev.* **32**, 643–671.
Benchimol, S., Fuks, A., Jothy, S., Beauchemin, N., Shirota, K. & Stanners, C. P. (1989). *Cell*, **57**, 327–334.
Bos, M. P., Grunert, F. & Belland, R. J. (1997). *Infect. Immun.* **65**, 2353–2361.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Duxbury, M. S., Ito, H., Benoit, E., Zinner, M. J., Ashley, S. W. & Whang, E. E. (2004). *Oncogene*, **23**, 5834–5842.
Dveksler, G. S., Pensiero, M. N., Cardellichio, C. B., Williams, R. K., Jiang, G.-S., Holmes, K. V. & Dieffenbach, C. W. (1991). *J. Virol.* **65**, 6881–6891.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Fedarovich, A., Tomberg, J., Nicholas, R. A. & Davies, C. (2006). *Acta Cryst.* D**62**, 971–979.
Gray-Owen, S. D. & Blumberg, R. S. (2006). *Nature Rev. Immunol.* **6**, 433–446.
Johnson, M. L., Correia, J. J., Yphantis, D. A. & Halvorson, H. R. (1981). *Biophys. J.* **36**, 575–588.
Korotkova, N., Yang, Y., Le Trong, I., Cota, E., Demeler, B., Marchant, J., Thomas, W. E., Stenkamp, R. E., Moseley, S. L. & Matthews, S. (2008). *Mol. Microbiol.* **67**, 420–434.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Kuroki, M., Abe, H., Imakiirei, T., Liao, S., Uchida, H., Yamauchi, Y.,

Oikawa, S. & Kuroki, M. (2001). *J. Leukoc. Biol.* **70**, 543–550.

Lawrence, M. C. & Colman, P. M. (1993). *J. Mol. Biol.* **234**, 946–950.

Leusch, H. G., Drzeniek, Z., Markos-Pusztai, Z. & Wagener, C. (1991). *Infect. Immun.* **59**, 2051–2057.

Litkouhi, B., Litkouhi, B., Fleming, E., Welch, W. R., Berkowitz, R. S., Birrer, M. J. & Mok, S. C. (2008). *Gynecol. Oncol.* **109**, 234–239.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Obrink, B. (2008). *Lung Cancer*, **60**, 309–312.

Oikawa, S., Inuzuka, C., Kuroki, M., Arakawa, F., Matsuoka, Y., Kosaki, G. & Nakazato, H. (1991). *J. Biol. Chem.* **266**, 7995–8001.

Oikawa, S., Inuzuka, C., Kuroki, M., Matsuoka, Y., Kosaki, G. & Nakazato, H. (1989). *Biochem. Biophys. Res. Commun.* **164**, 39–45.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Pils, S., Gerrard, D. T., Meyer, A. & Hauck, C. R. (2008). *Int. J. Med. Microbiol.* **298**, 553–560.

Roda, G., Jianyu, X., Park, M. S., DeMarte, L., Hovhannisyan, Z., Couri, R., Stanners, C. P., Yeretssian, G. & Mayer, L. (2014). *Mucosal Immunol.* **7**, 615–624.

Schölzel, S., Zimmermann, W., Schwarzkopf, G., Grunert, F., Rogaczewski, B. & Thompson, J. (2000). *Am. J. Pathol.* **156**, 595–605.

Singer, B. B., Opp, L., Heinrich, A., Schreiber, F., Binding-Liermann, R., Berrocal-Almanza, L. C., Heyl, K. A., Müller, M. M., Weimann, A., Zweigner, J. & Slevogt, H. (2014). *PLoS One*, **9**, e94106.

Singer, B. B., Scheffrahn, I., Kammerer, R., Suttorp, N., Ergun, S. & Slevogt, H. (2010). *PLoS One*, **5**, e8747.

Skubitz, K. M. & Skubitz, A. P. (2008). *J. Transl. Med.* **6**, 78.

Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.

Streichert, T., Ebrahimnejad, A., Ganzer, S., Flayeh, R., Wagener, C. & Brümmer, J. (2001). *Biochem. Biophys. Res. Commun.* **289**, 191–197.

Taheri, M., Saragovi, U., Fuks, A., Makkerh, J., Mort, J. & Stanners, C. P. (2000). *J. Biol. Chem.* **275**, 26935–26943.

Tchoupa, A. K., Schuhmacher, T. & Hauck, C. R. (2014). *Cell Commun. Signal.* **2014**, 12–27.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.

Virji, M., Evans, D., Griffith, J., Hill, D., Serino, L., Hadfield, A. & Watt, S. M. (2000). *Mol. Microbiol.* **36**, 784–795.

Watt, S. M., Teixeira, A. M., Zhou, G.-Q., Doyonnas, R., Zhang, Y., Grunert, F., Blumberg, R. S., Kuroki, M., Skubitz, K. M. & Bates, P. A. (2001). *Blood*, **98**, 1469–1479.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Zheng, J., Miller, K. K., Yang, T., Hildebrand, M. S., Shearer, A. E., DeLuca, A. P., Scheetz, T. E., Drummond, J., Scherer, S. E., Legan, P. K., Goodyear, R. J., Richardson, G. P., Cheatham, M. A., Smith, R. J. & Dallos, P. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 4218–4223.