



Phosphorylation of human CEACAM1-LF by PKA and GSK3 β promotes its interaction with β -catenin

Received for publication, July 18, 2021, and in revised form, October 5, 2021. Published, Papers in Press, October 14, 2021.
<https://doi.org/10.1016/j.jbc.2021.101305>

Weidong Hu¹, Karine Bagramyan¹, Supriyo Bhattacharya², Teresa Hong¹, Alonso Tapia¹, Patty Wong¹, Markus Kalkum¹, and John E. Shively^{1,*}

From the ¹Department of Molecular Imaging and Therapy, and ²Department of Computational and Quantitative Medicine, Beckman Research Institute of City of Hope, Duarte, California, USA

Edited by Peter Cresswell

CEACAM1-LF, a homotypic cell adhesion molecule, transduces intracellular signals *via* a 72 amino acid cytoplasmic domain that contains two immunoreceptor tyrosine-based inhibitory motifs (ITIMs) and a binding site for β -catenin. Phosphorylation of Ser503 by PKC in rodent CEACAM1 was shown to affect bile acid transport or hepatosteatosis *via* the level of ITIM phosphorylation, but the phosphorylation of the equivalent residue in human CEACAM1 (Ser508) was unclear. Here we studied this analogous phosphorylation by NMR analysis of the ¹⁵N labeled cytoplasmic domain peptide. Incubation with a variety of Ser/Thr kinases revealed phosphorylation of Ser508 by GSK3 β but not by PKC. The lack of phosphorylation by PKC is likely due to evolutionary sequence changes between the rodent and human genes. Phosphorylation site assignment by mass spectrometry and NMR revealed phosphorylation of Ser472, Ser461 and Ser512 by PKA, of which Ser512 is part of a conserved consensus site for GSK3 β binding. We showed here that only after phosphorylation of Ser512 by PKA was GSK3 β able to phosphorylate Ser508. Phosphorylation of Ser512 by PKA promoted a tight association with the armadillo repeat domain of β -catenin at an extended region spanning the ITIMs of CEACAM1. The kinetics of phosphorylation of the ITIMs by Src, as well dephosphorylation by SHP2, were affected by the presence of Ser508/512 phosphorylation, suggesting that PKA and GSK3 β may regulate the signal transduction activity of human CEACAM1-LF. The interaction of CEACAM1-LF with β -catenin promoted by PKA is suggestive of a tight association between the two ITIMs of CEACAM1-LF.

CEACAM1 is a homotypic cell adhesion molecule expressed in many cell types with a variety of signal transduction properties conveyed by its cytoplasmic domains (1, 2). Due to alternative mRNA splicing, the cytoplasmic domain may be expressed as either long (CEACAM1-LF, about 72 amino acids) or short (CEACAM1-SF, about 12 amino acids) isoforms (3, 4). The cytoplasmic domain of CEACAM1-LF contains two immunoreceptor tyrosine-based inhibitory motifs (ITIMs; Fig. 1) that when phosphorylated by Src kinases (5)

recruit the inhibitory phosphatases SHP1 (6, 7) or SHP2 (8), limiting signaling by other cell tyrosine phosphorylated surface receptors (9). Important biological functions of CEACAM1-LF include a role in the immune response in lymphocytes (1) and insulin clearance in the liver (10). A highly conserved sequence spanning residues 462 to 471 in rodents and 467 to 476 in man (Fig. 1) has been shown to bind to β -catenin (11–14), a key signaling molecule involved in cell adhesion (15) and gene transcription (16). The CEACAM1 knockout mouse has major effects on tumor progression (17), immune responses (1) and insulin signaling (10) with a remarkable defect in lipid storage in male mice (18, 19). The cytoplasmic domain of CEACAM1-SF binds G-actin (20) and Annexin A2 (21, 22) *via* the key residues Phe454 and Thr457 in human CEACAM1, where both residues are shared in the proximal sequences of both isoforms (Fig. 1). Re-expression of CEACAM1-SF reverts breast cancer cells to a normal phenotype (23, 24). Expression of the two isoforms varies among cell types, with CEACAM1-LF expressed mainly in lymphocytes (1), CEACAM1-SF mainly in epithelial cells (25), and both isoforms in hepatocytes (26, 27).

Due to the extensive sequence homology of CEACAM1 cytoplasmic domains among mammalian species (28), it has been assumed that signaling the phenotypes deduced from rodents applies to humans. However, a comparison of the rodent and human sequences reveals interesting evolutionary changes that challenge this assumption (Fig. 1). For example, the end of the transmembrane and start of the cytoplasmic domains are in doubt between rodent and human CEACAM1. One of the double basic residues, Arg450-Lys451 in the rodent sequences, a classic halt sequence at the end of a transmembrane domain, is reduced to a single basic residue Lys456 in humans. Furthermore, residue Phe454 that precedes Lys456 and was shown to bind G-actin and annexin A2 in humans is missing in rodents. Although many Ser/Thr residues in CEACAM1 are potential phosphorylation sites, the adjoining sequences that often determine kinase specificity are notably different. This is especially true of Ser503 in rat that has proximal double basic residues Lys499 and Arg500 preceding Ser503, while mouse has a single basic residue, and human none at this location. The importance of this difference becomes clear when one realizes that Ser503 was shown to be a

* For correspondence: John E. Shively, jshively@coh.org.

Phosphorylation of CEACAM1 by PKA and GSK3 β

target for PKC phosphorylation in rat and its mutation to Ala abrogated bile acid uptake in a heterologous cell reporter (29). Furthermore, the null mutation Ser503Ala in rat (29) and in mouse (30–32) was shown to affect the phosphorylation of the ITIM residue Tyr488.

A transgenic mouse expressing the null mutation Ser502Ala in *Ceacam1* that developed insulin resistance and fatty liver disease (33) had a similar phenotype to the *Ceacam1* knockout mouse (34). Furthermore, the null mutation Ser503Ala in murine *Ceacam1* was shown to affect insulin signaling (31). Thus, it was thought that Ser503 was a key determinant in the signal transduction of the cytoplasmic domain of CEACAM1-LF. However, the transgenic mouse model can be questioned on the basis of coexpression of wild-type CEACAM1. These questions become important when trying to apply rodent mutation studies to man. Based on the sequence differences surrounding the rodent Ser503 and human equivalent Ser508 residue (Fig. 1), we speculate that the kinase responsible for phosphorylation of human Ser508 may not be PKC. In fact, the presence of conserved Ser512 four residues downstream of Ser508 conforms to a GSK3 β phosphorylation consensus site (35), that if verified, would strongly link the function of CEACAM1 to insulin-mediated glucose and lipid regulation, since the action of insulin on its receptor results in the phosphorylation of Akt, which phosphorylates GSK3 β , ultimately inhibiting glycogen synthase (35) and affecting both glucose and lipid regulation (36).

A further source of confusion is the assertion that CEACAM1 is a bile acid transporter and that the null mutation Ser503Ala in rat CEACAM1-LF was able to abrogate this function (29). This was a tempting conclusion, since the major biliary protein BGP1 (37) was later shown to be identical to CEACAM1 (38). However, later studies unambiguously demonstrated that bile acid transporters NTCP, BSEP, ASBT, and OST α/β were responsible for the import and export of bile acids (39). Thus, the functional status of CEACAM1 in the bile remains unclear.

Although mouse models for protein function have many correlates in man, many cautionary studies demonstrate that evolutionary changes in amino acid sequences can dictate changes in function (40). We believe this may be the case for Ser508 in human CEACAM1-LF. As a first step in approaching the problem, we tested the activity of a variety of kinases on the cytoplasmic domain of human CEACAM1-LF expressed in *E. coli* with and without ^{15}N -labeling, the latter allowing

assignment of residues by 2D NMR, as well as phosphorylation kinetic studies. Since we had previously predicted that a phosphorylated form of CEACAM1-LF would bind β -catenin more strongly than unphosphorylated CEACAM1-LF (11), we also performed binding studies of β -catenin to the peptide before and after phosphorylation and built a model of their interaction based on the NMR data and molecular modeling.

Results

Screening of Ser/Thr kinases for CEACAM1-LF cytoplasmic domain phosphorylation using 2D ^1H - ^{15}N HSQC and 1D ^1H NMR

The backbone assignment of the 74 amino acid cytoplasmic domain of human CEACAM1-LF was performed using 2D and 3D triple resonance experiments carried out on ^{15}N , ^{13}C -labeled CEACAM1-LF peptide at pH 5.5 as described in the Experimental procedures. The backbone chemical shift analysis using TALOS-N (41) showed the CEACAM1 peptide is unstructured and very mobile (Fig. S1). Initial assignments acquired at pH 5.5 from the ^1H - ^{15}N HSQC spectra are shown in Figure 2A. The assignments at pH 6.6 (Fig. 2B) were obtained by a series of titrations from pH 5.5 to 6.6 and were used to monitor phosphorylation of ^{15}N -CEACAM1 by a variety of kinases through the comparison of the spectrum in the absence and presence of the kinases. The tested Ser/Thr and Tyr kinase or mixture of kinases is listed in Table S1. Among the tested kinases, PKA, PAK1, and Src were able to phosphorylate the peptide, while PKC ϵ and other tested Ser/Thr kinases did not phosphorylate the peptide under the same conditions. After phosphorylation, the amide proton chemical shift of Ser/Thr residues moved down field by 0.5 to 1.5 ppm (42). These chemical shift changes are mainly caused by the intraresidue hydrogen bonds formation between amide protons and the phosphate group (43), which make the phosphorylated amide group well resolved from other amide proton chemical shifts in an unstructured peptide, thus enabling the longitudinal study of phosphorylation through 1D ^1H NMR. The phosphorylation of unlabeled CEACAM1-LF peptide by PKA followed by GSK3 β as monitored by 1D ^1H NMR is shown in Figure 3. Compared with the control spectrum in the absence of PKA (Fig. 3A), three strong (b, c, d) and one weak (a) peaks appeared between 8.7 and 9.1 ppm after phosphorylation by PKA (Fig. 3B). This indicated at least three Ser/Thr residues in the peptide were phosphorylated by

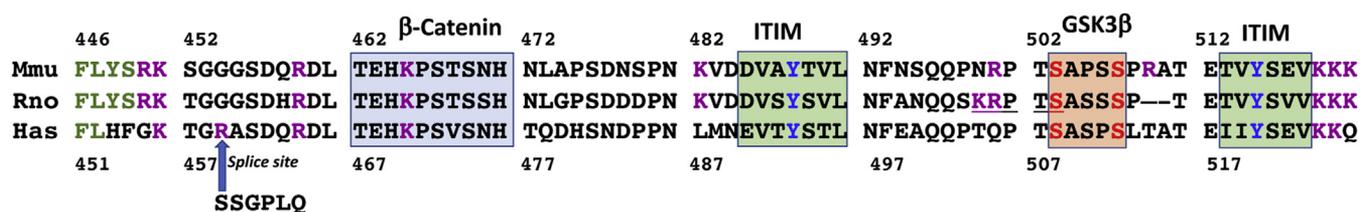


Figure 1. Sequence comparison of murine, rat, and human CEACAM1-LF cytoplasmic domains. Presumed residues within the transmembrane domain are in green, basic residues are in magenta, tyrosines in blue, and predicted serines in the GSK3 β consensus site in red. The conserved β -catenin site is boxed in blue, ITIM sites in green, and GSK3 β site in orange. The location of the mRNA splice site leading to formation of CEACAM1-SF in human is indicated with an arrow along with the amino acid sequence resulting from the splice site.

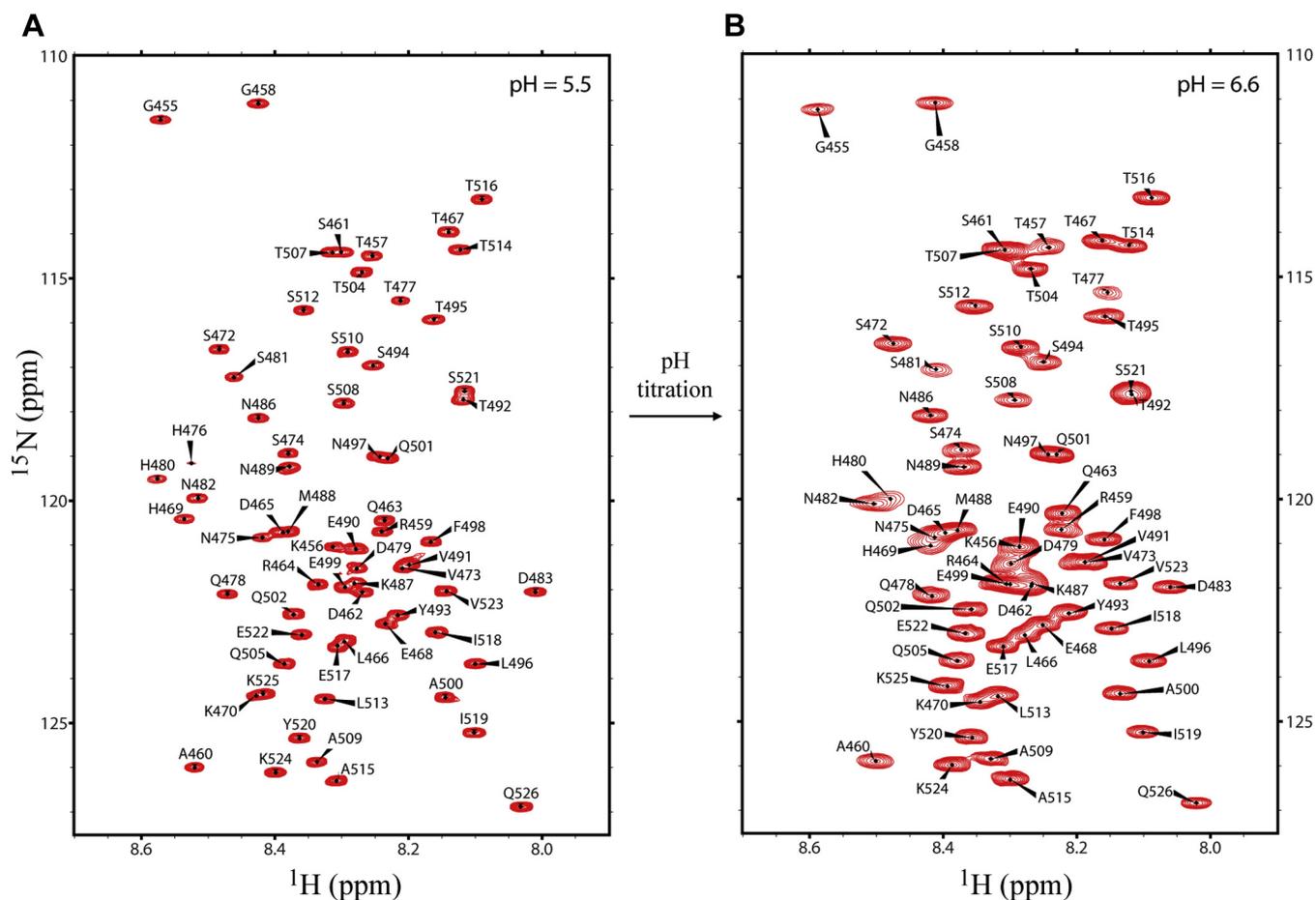


Figure 2. 2D NMR spectra of ^{15}N labeled CEACAM1-LF cytoplasmic domain peptide. A, pH 5.5. B, pH 6.6. The assignment of CEACAM1-LF peptide at pH 6.6 was derived from tracking peak position changes versus a series of pH values of 5.5, 5.7, 5.9, 6.1, 6.2, 6.4, 6.6, and 6.7. Residue numbers are according to Figure 1.

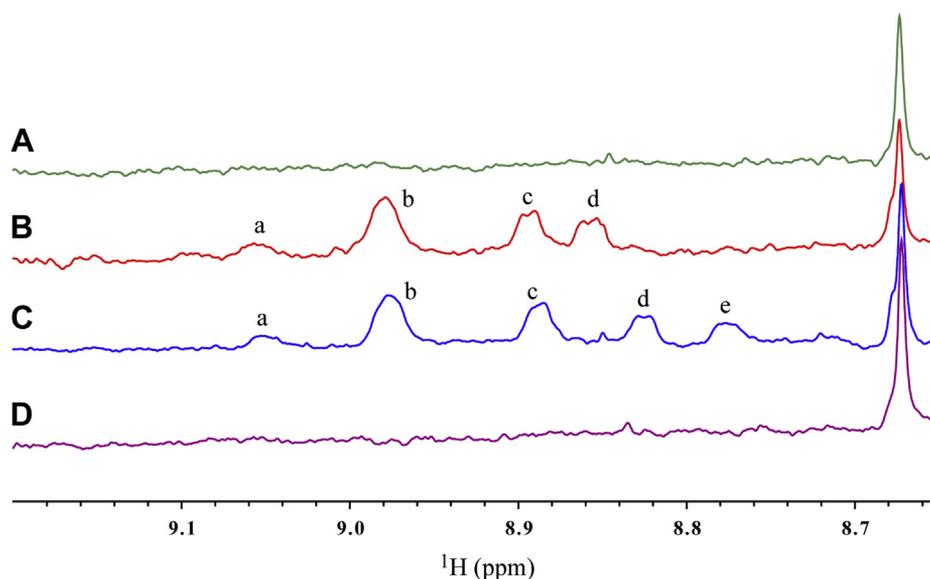


Figure 3. Monitoring kinase reactions on CEACAM1-LF cytoplasmic domain peptide by ^1H NMR. A, the down field amide region of a ^1H NMR spectrum acquired immediately after addition of PKA to the CEACAM1-LF. B, the ^1H NMR spectrum after 5 days at room temperature. Four new peaks appear in the monitored region due to phosphorylation of either Thr or Ser residues. C, the ^1H NMR after 2 days of addition of GSK3 β to the sample in (B) after denaturing the PKA at 60 $^\circ\text{C}$ for 30 min. See [Experimental procedures](#) for more details. New peak "e" appeared due to additional Thr or Ser phosphorylation by GSK3 β . The peak position of "d" moved compared with that in spectrum B likely due to the close proximity of the new phosphorylated residue "e". D, the ^1H NMR of CEACAM1-LF at 2 days after addition of the GSK3 β without pretreatment of PKA.

Phosphorylation of CEACAM1 by PKA and GSK3 β

PKA. The sample was then heated at 60 °C for 30 min to denature the PKA, after which GSK3 β was added, resulting in an up-field shift of peak **d** and the appearance of new peak **e** (Fig. 3C). This result suggested that GSK3 β phosphorylated an additional Ser or Thr in the peptide, and that this residue was proximal to peak **d**. The control spectrum in which the peptide was treated with GSK3 β prior to PKA demonstrated that GSK3 β was unable to phosphorylate the native peptide (Fig. 3D).

Identification of phosphorylation sites by mass spectrometry

Following 1D NMR analysis of unlabeled CEACAM1-LF peptide before and after kinase treatment, the peptides were analyzed by high-resolution ESI/MS and MALDI/MS (Fig. 4). The results are consistent with the presence of three major phosphorylated sites by PKA and a fourth site after the addition of GSK3 β . To identify the precise sites of phosphorylation, the peptide samples were digested with endo Lys C and analyzed by multistage LC/MS. The analysis covered the entire sequence except for the first four amino acids (Fig. S2) and identified Ser512 as one of the sites of PKA phosphorylation and Ser508 as the site of GSK3 β phosphorylation after

treatment with PKA (Fig. 5 and Tables S3–S7). The two additional PKA sites were identified as Ser461 and Ser 472 (data not shown).

Identification of PKA phosphorylation sites in CEACAM1-LF peptide by 2D NMR

When Ser/Thr residues are phosphorylated in an unstructured peptide, their proton chemical shifts are increased significantly, as well as the chemical shift perturbation of nearby residues, but to a lesser extent (43). The chemical shift changes in 2D spectra of the ^{15}N -CEACAM1-LF peptide were monitored over time of treatment with PKA, and three of those time points are presented in (Fig. 6). The results demonstrate that Ser 472 is phosphorylated first (Fig. 6A), followed by Ser461 and Ser512 (Fig. 6, B and C). As expected, phosphorylation of these three residues caused minor perturbations in adjacent residues. For example, phosphorylation of Ser472 perturbed residues H469, K470, S474, T477, and N475; phosphorylation of Ser461 perturbed T457, G458, A460 and Q463; and phosphorylation of S512 perturbed S508, L513, T514 and A515. The assignments of peak positions in Figure 6C (red spectrum) are based on agreement of their intensity change rate

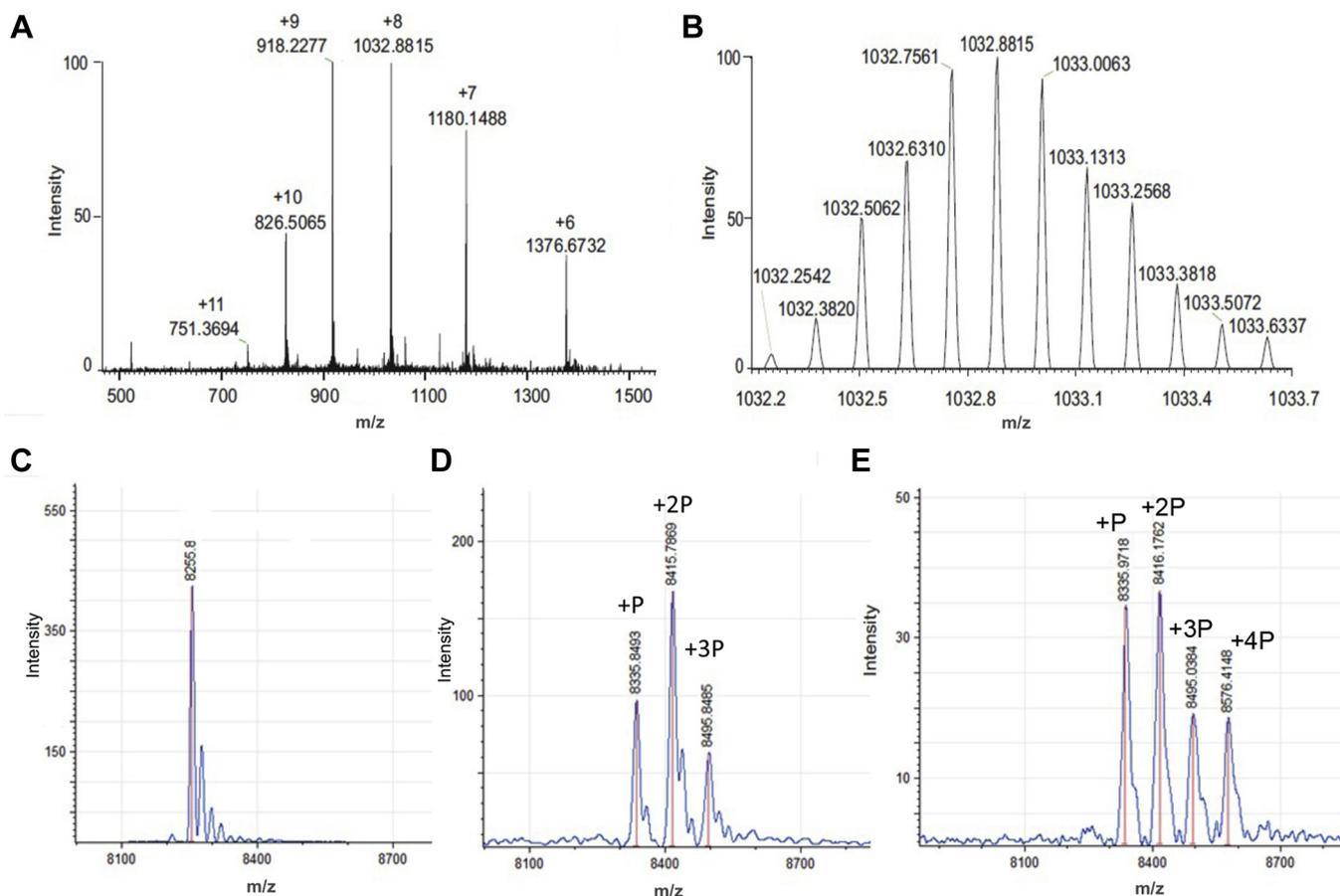


Figure 4. Mass spectrometric analysis of CEACAM1-LF cytoplasmic domain peptide before and after treatment with PKA or PKA plus GSK3 β . **A**, ESI orbitrap MS of the intact CEACAM1-LF peptide showing charge states. **B**, high-resolution analysis of the $z = 8+$ charge state. The deduced average mass was 8249.977 Da (expected 8249.965 Da). The deduced average mass was 8255.14 Da (expected 8254.94 Da). **C**, MALDI-MS analysis of the intact peptide, expected $M + H^+$ at m/z 8255 (observed m/z 8256). **D**, MALDI-MS analysis after treatment of the peptide with PKA. The intact peptide was modified by the addition of three phosphates (80 mass units per phosphate, $+n\text{P}$). **E**, MALDI/MS analysis after treatment with PKA plus GSK3 β . The intact peptide was modified by the addition of up to four phosphates ($+4\text{P}$).

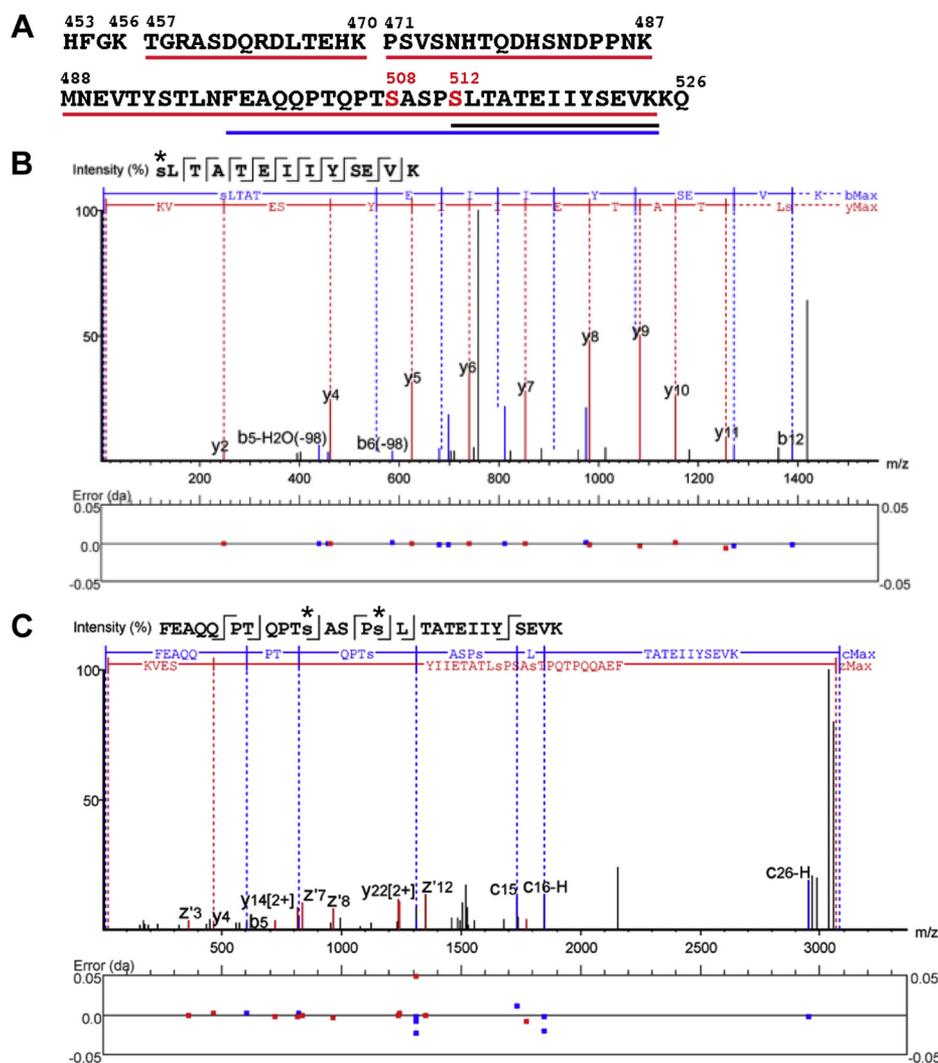


Figure 5. LC-MS/MS analysis of CEACAM1-LF cytoplasmic domain peptide. A, mass spectrometric evidence was obtained for the entire sequence (underlined in red) after digestion with endoproteinase-LysC with exception of the first four amino acids. Phosphorylated Ser508 and 512 are shown in red. B, MS/MS spectra of a phosphopeptide (underlined in black) after treatment with PKA. C, MS/MS spectra of phosphopeptide (underlined in blue) after treatment with PKA + GSK3 β . The phosphorylated serine residues are marked with a lower case “s” and an asterisk. Fragment ion assignments and mass differences from calculated values are shown in the mass error maps underneath each spectrum (B and C).

relative to the three phosphorylated Ser residues, and their peak position movements relative to their corresponding initial peak positions (blue spectrum). To further verify the phosphorylation site of S512, a 37 amino acid C-terminal peptide (E490-Q526) of CEACAM1-LF was synthesized in which S508 and S512 were ^{15}N -labeled. When this peptide was treated with PKA, the cross peak position of phosphorylated S512 was identical to that of phosphorylated S512 in the full-length ^{15}N -labeled CEACAM1-LF peptide (Fig. S3A).

Identification of GSK3 β phosphorylation site in CEACAM1-LF peptide after treatment with PKA by 2D NMR

After Ser512 in the ^{15}N -peptide was fully phosphorylated by PKA, as confirmed by the 2D spectra, the sample was heated at 60 °C for 30 min to denature the PKA and separated from the phosphorylated CEACAM1-LF peptide as described in Experimental procedures. The PKA phosphorylated peptide was then treated with GSK3 β (Fig. 7A). Ser508 was identified

as a new site of phosphorylation along with minor perturbations in adjacent residues S510, pS512 (phosphorylated Ser512), and L513. The same 37 amino acid peptide used in Fig. S3A was used to further confirm that Ser508 was phosphorylated by GSK3 β after phosphorylation of Ser512 by PKA. When the 2D spectrum of this 37 amino acid peptide, sequentially treated with PKA and GSK3 β , was overlaid with the spectrum of the full-length ^{15}N -CEACAM1-LF peptide, also phosphorylated by PKA and GSK3 β , the peak position of phosphorylated Ser508 was identical (Fig. S3B).

Phosphorylation of tyrosines in CEACAM1-LF peptide by Src before and after phosphorylation of Ser508

Since it was reported that phosphorylation of Ser503 in rodent CEACAM1-LF affected phosphorylation of Tyr488 (29), we tested the effect of PKA plus GSK3 β on the phosphorylation of tyrosines Tyr493 and Tyr520 in the ^{15}N -peptide of human CEACAM1-LF. Since Src is the relevant tyrosine

Phosphorylation of CEACAM1 by PKA and GSK3 β

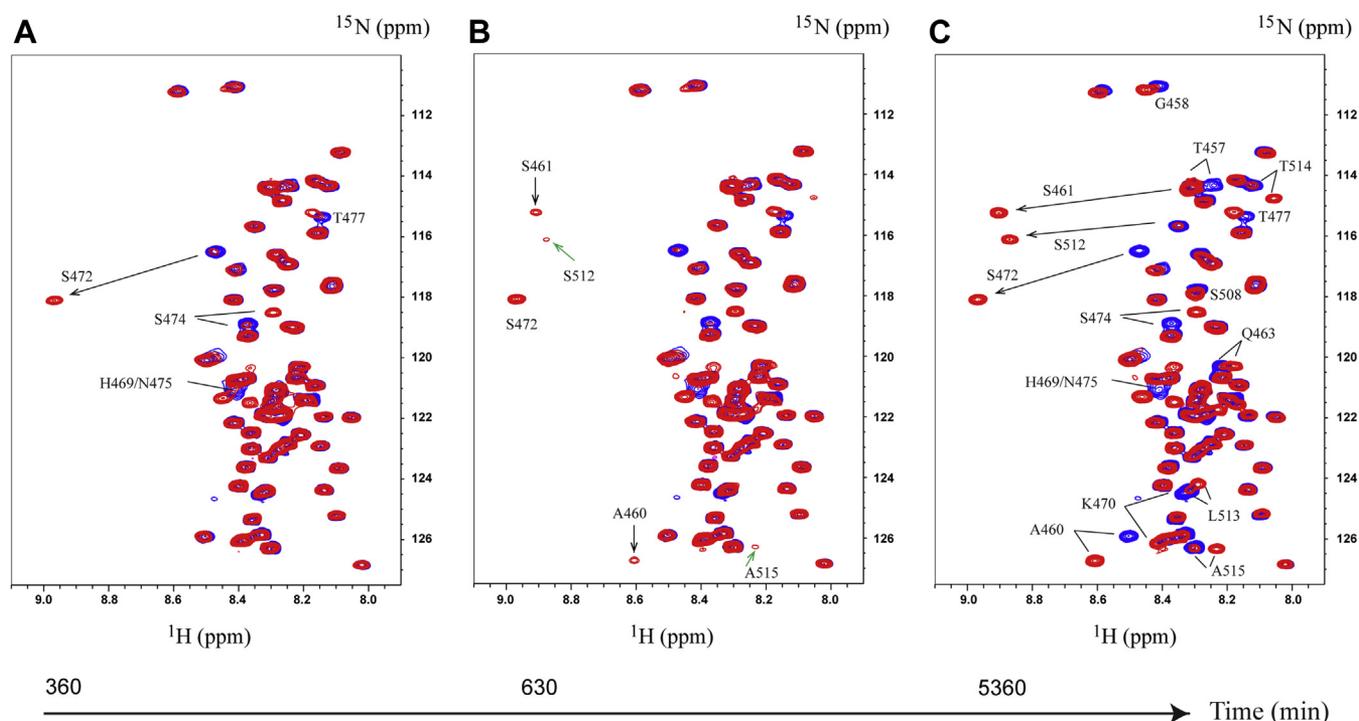


Figure 6. Longitudinal ^1H - ^{15}N spectra of CEACAM1-LF cytoplasmic domain peptide treated with PKA. The spectra in *blue* (A–C) were acquired on CEACAM1-LF peptide without PKA. The spectra in *red* were acquired at three time points (A–C) after addition of PKA. *A*, after 360 min, S472 was almost completely phosphorylated. A few residues close to S472 with chemical shift perturbation were labeled. *B*, phosphorylation peaks of S461 and S512 together with nearby residues were found at new positions as indicated by *black* and *green* arrows, respectively. *C*, phosphorylation of S472 and S461 was complete at the indicated time, and majority of S512 was phosphorylated. Some residues close to the three phosphorylated serines with notable chemical shift perturbation are labeled.

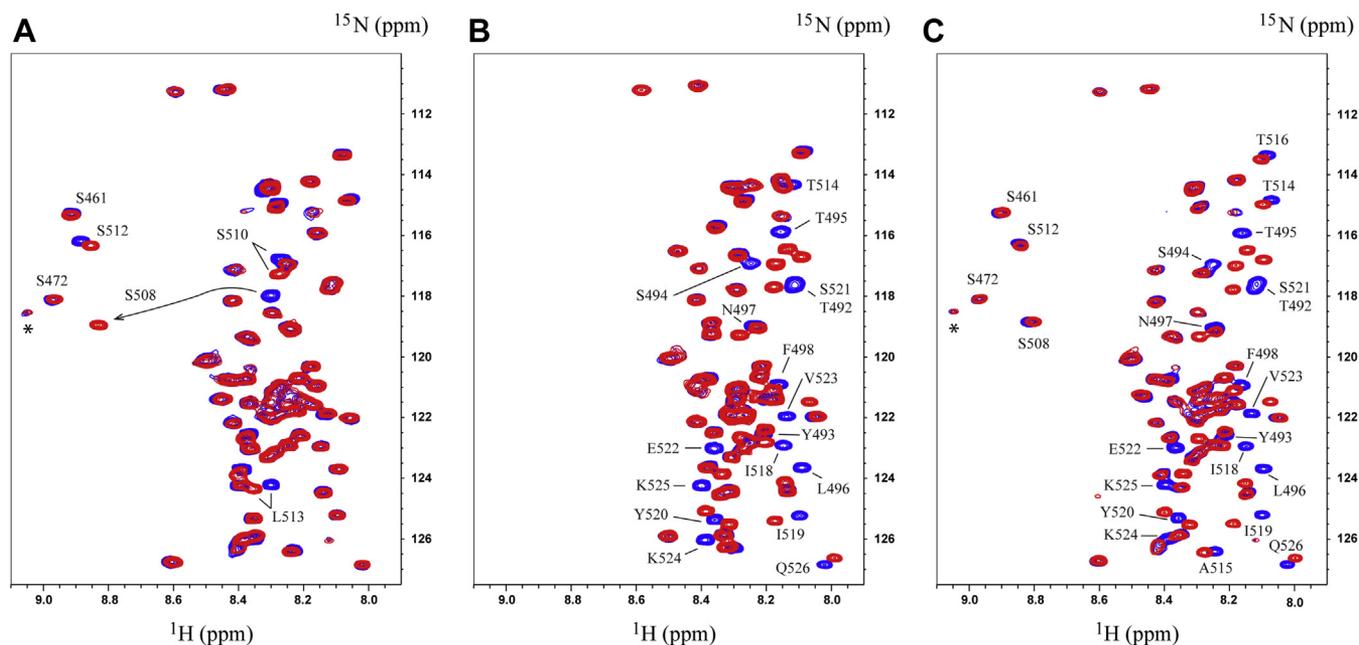


Figure 7. Phosphorylation of Ser508 by GSK3 β after treatment with PKA and phosphorylation Tyr by Src. *A*, spectra overlay of phosphorylated CEACAM1-LF peptide by PKA (in *blue*) and further phosphorylated by GSK3 β (in *red*). The residues close to S508 with notable chemical shift perturbations are labeled. The peak labeled with “*” is from an unassigned residual phosphorylated residue by PKA. *B*, spectra overlay of CEACAM1-LF (in *blue*) and phosphorylated CEACAM1-LF by Src (in *red*). The residues close to Tyr493 and Tyr520 with significant chemical shift perturbations caused by the two Tyr phosphorylation are labeled. *C*, spectra overlay of phosphorylated CEACAM1-LF by PKA and GSK3 β (in *blue*) and further phosphorylated by Src (in *red*). The residues close to Tyr493 and Tyr520 with significant chemical shift perturbations caused by the two Tyr phosphorylation are labeled. The peak labeled with “*” is from an unassigned residual phosphorylated residue by PKA.

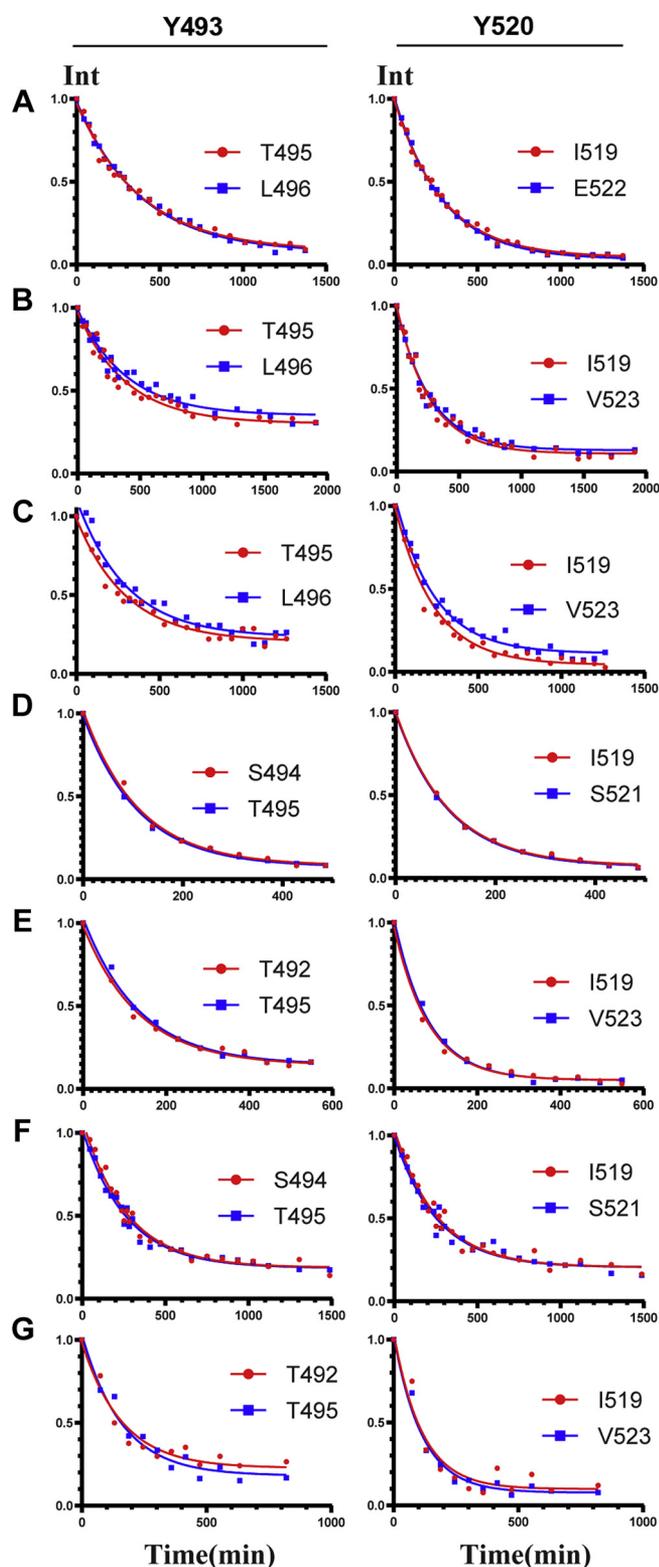


Figure 8. Kinetics of phosphorylation of ITIM tyrosines in CEACAM1-LF peptide by Src and their dephosphorylation by SHP2. A–C, kinetics of phosphorylation of Y493 and Y520 after treatment of the CEACAM1-LF peptide with Src (A) or PKA and GSK3 β then Src (B) or PKA and GSK3 β then Src in the presence of β -catenin (C). D–G, SHP2 dephosphorylation of peptide phosphorylated by Src (D) or PKA and GSK3 β then Src (E) or Src in the presence of β -catenin (F) or PKA and GSK3 β then Src in the presence of β -catenin (G). Peak intensities are monitored from the residues shown in insets.

kinase for CEACAM1-LF (5), the peptide was treated with Src before (Fig. 7B) and after treatment with PKA plus GSK3 β (Fig. 7C). The native peptide was readily phosphorylated by Src on residues Y493 and Y520, along with perturbation in their adjacent residues (T492, S494, T495, L496, N497, F498; T514, A515, T516, I518, I519, S521, E522, V523, K524, K525, Q526). The two tyrosines were also phosphorylated by Src when treated with PKA plus GSK3 β with similar chemical shift perturbations in the residues adjacent to Y493 and Y520 as observed in the native peptide (Fig. 7B).

Kinetics of Src phosphorylation and SHP2 dephosphorylation of Tyr493 and Tyr520 and effect of PKA and GSK3 β phosphorylation

Since CEACAM1-LF has two ITIM tyrosines that are phosphorylated by Src (5), we were interested if the kinetics of their phosphorylation were similar and if they were affected by phosphorylation of Ser508/512, located between the two ITIMs. Cross peak intensity changes in residues T495 and L496 adjacent to Tyr493 and residues I519 and V522 adjacent to Tyr520 were chosen to monitor the kinetics of Src phosphorylation of the native peptide (Fig. 8A). The ratio of K_{app} between the two residues indicates a preference of Src for distal Tyr520 over proximal Tyr493 by a factor of 1.3 that is not significantly affected by pSer512 (Fig. 8B and Table 1). In addition, the kinetics of Src phosphorylation of the peptide prephosphorylated with PKA plus GSK3 β were not affected by the presence of β -catenin (Fig. 8C and Table 1), suggesting that the preference of Src for Tyr520 was intrinsic to the surrounding sequence/accessibility of this residue. Although Src has an apparent preference in terms of rates of phosphorylation of the two tyrosines, it appears that the dephosphorylation rates of these two tyrosines are comparable when treated with SHP2 (Fig. 8D and Table 1). Since SHP2 has two SH2-binding domains, it was possible that it would bind to one phosphotyrosine preferentially and dephosphorylate the other site. However, the lack of discrimination between the dephosphorylation of the two phospho-tyrosines was dramatically altered when the study was repeated on CEACAM1-LF peptide pretreated with PKA plus GSK3 β (Fig. 8E and Table 1). This result suggests that SH2 binding of SHP2 may play a role in CEACAM1-LF that has been phosphorylated on Ser508/512 located between the two ITIMs. Notably, the addition of β -catenin to the Src phosphorylated peptide before or after PKA plus GSK3 β treatment had no effect on the rates of dephosphorylation of the two tyrosines by SHP2 (Fig. 8, F and G and Table 1).

Binding of β -catenin to CEACAM1-LF peptide before and after treatment with PKA

Since we had previously shown that the cytoplasmic domain of human CEACAM1-LF could bind β -catenin (11), we were interested if phosphorylation by PKA plus GSK3 β would affect the binding. First, we incubated the untreated peptide with β -catenin and found no changes in chemical shifts over time,

Phosphorylation of CEACAM1 by PKA and GSK3 β

Table 1

Kinetic analysis of phosphorylation of Tyr493 and Tyr520 on CEACAM1-LF cytoplasmic domain by Src and their dephosphorylation by SHP2 in the absence or presence of β -catenin

Treatments ^a	K_{app} (min ⁻¹) \times 1000		K_{app} ratio
	Tyr493	Tyr520	Tyr520/Tyr493
Src on CEACAM1	2.59 \pm 0.22	3.45 \pm 0.15	1.34 \pm 0.094
Src on PKA/GSK3 β -CEACAM1	2.55 \pm 0.36	3.72 \pm 0.35	1.46 \pm 0.17
Src on PKA/GSK3 β -CEACAM1 + β -catenin	3.32 \pm 0.46	4.25 \pm 0.39	1.28 \pm 0.17
SHP2 on Src-CEACAM1	8.99 \pm 0.86	9.38 \pm 0.59	1.04 \pm 0.11
SHP2 on PKA/GSK3 β /Src-CEACAM1	7.43 \pm 0.83	12.09 \pm 1.08	1.63 \pm 0.14
SHP2 on Src-CEACAM1 + β -catenin	3.76 \pm 0.36	3.89 \pm 0.46	1.03 \pm 0.15
SHP2 on PKA/GSK3 β /Src-CEACAM1 + β -catenin	6.24 \pm 1.35	8.86 \pm 1.69	1.42 \pm 0.29

^a PKA/GSK3 β -CEACAM1 denotes pretreatment of the cytoplasmic domain peptide with PKA and then GSK3 β .

suggesting minimal or no binding (data not shown). However, this result was not unexpected since our previous modeling study was predicated on the requirement for phosphorylated residues in the proximal region of the peptide to bind β -catenin (11). Since PKA was able to phosphorylate three Ser residues spanning the proximal to distal regions of the peptide (Ser461, Ser472, and Ser512), we next incubated the unphosphorylated peptide with β -catenin (1.7 M equivalents) and added PKA to determine if the above determined phosphorylation sites would induce binding. The time course study shows in the 2D spectra of the peptide plus β -catenin with the addition of PKA (Fig. 9, A–C). In Figure 9, A–C, the spectra in blue are the mixture of CEACAM1 and β -catenin before addition of PKA, and the spectra in red are acquired at 49, 69 and 74 h, respectively, after addition of PKA. The phosphorylated cross peaks of S461, S472, and S512 by PKA grow with time, with the highest phosphorylation states for S472, followed by S461 and S512. The phosphorylation rate order is the same as that found in the absence of β -catenin (Fig. 6C). Over time, S472 was completely phosphorylated, there was no further increase in pS461, and the pS512 peak disappeared. In addition, the intensity of residues at the C-terminus decreased (Fig. 9, B and C). Because of the relatively high molecular weight of β -catenin (about 86 kDa) versus the ¹⁵N-labeled phosphorylated peptide, the peak intensities of the peptide residues tightly associated with β -catenin were expected to gradually diminish over the course of the reaction. The intensity changes of cross peaks are shown as an overlay of spectrum in red (Fig. 9C) versus fully phosphorylated CEACAM1 by PKA in the absence of β -catenin along with assignments in the spectrum in blue (Fig. 9D). A plot of the peak intensity ratio versus residue number is shown in Figure 10. The results suggest that tightest binding of β -catenin to the peptide occurred between the two ITIMs in CEACAM1. Since Ser512 is the only PKA phosphorylated residue in this region, this result suggests that pSer512 is a key residue for β -catenin binding and that this binding may affect the phosphorylation status of both tyrosines. It should be noted that although there is a substantial reduction in apparent signal intensity of Val473 adjacent to Ser472 and Arg464 near Ser461, their lower intensities are likely due to the larger signal reductions in V491 and E499 due to their overlapping cross peaks. The overall lack of reduction in signal intensity outside the ITIMs suggests that the segment between the two ITIMs plays a critical role in β -catenin binding.

When the ¹⁵N-peptide was prephosphorylated with PKA, or PKA plus GSK3 β followed by the addition of β -catenin, no changes in signal intensities were observed, in contrast to the results observed for the coincubated peptide with β -catenin plus PKA. Since PKA has been shown to phosphorylate β -catenin increasing its binding to substrates (44–46), we analyzed the PKA-treated β -catenin by mass spectrometry and found evidence of phosphorylation at previously reported Ser552 (44) and at the additional residue Ser179 (Fig. S4). Since both of these residues lie within the armadillo repeats of β -catenin, these results suggest that PKA phosphorylation of both the cytoplasmic domain of CEACAM1-LF and β -catenin is required for their association.

An in silico study of the binding of PKA phosphorylated β -catenin and CEACAM1-LF peptide

The above results indicated that PKA phosphorylation of both β -catenin and CEACAM1-LF was required for their interaction. Phosphorylated Ser552 on β -catenin is located on an extensive loop on armadillo repeat 10, defining a region that juts up from the otherwise flat structure (47), while phosphorylated Ser179 is located at the N-terminal armadillo repeat 1. In order to build a model of the complex CEACAM1-LF and β -catenin, a sequence alignment was performed with known structures (Fig. S5), along with pSer552 and pSer179 on β -catenin and pSer512 on CEACAM1-LF as anchors, followed by energy minimization of the resulting structure (Fig. 11). The proposed structure predicts interactions of pSer552 on β -catenin with Arg464 on CEACAM1-LF, and pSer179 on β -catenin with Lys525 on CEACAM1-LF. Phosphorylated Ser512 on CEACAM1-LF is buried deep in the groove of β -catenin (inset Fig. 11), in agreement with its requirement for binding to β -catenin, while Ser508 is exposed, allowing its phosphorylation by GSK3 β even after binding to β -catenin. The main interactions of CEACAM1-LF with β -catenin are between the ITIMs as predicted by the NMR studies. Future studies that would test the model include X-ray structural analysis of the phosphorylated complex along with mutational analysis of β -catenin.

Discussion

Among the biological functions ascribed to CEACAM1, signaling from the long cytoplasmic isoform CEACAM1-LF has attracted the most interest due to the presence of two

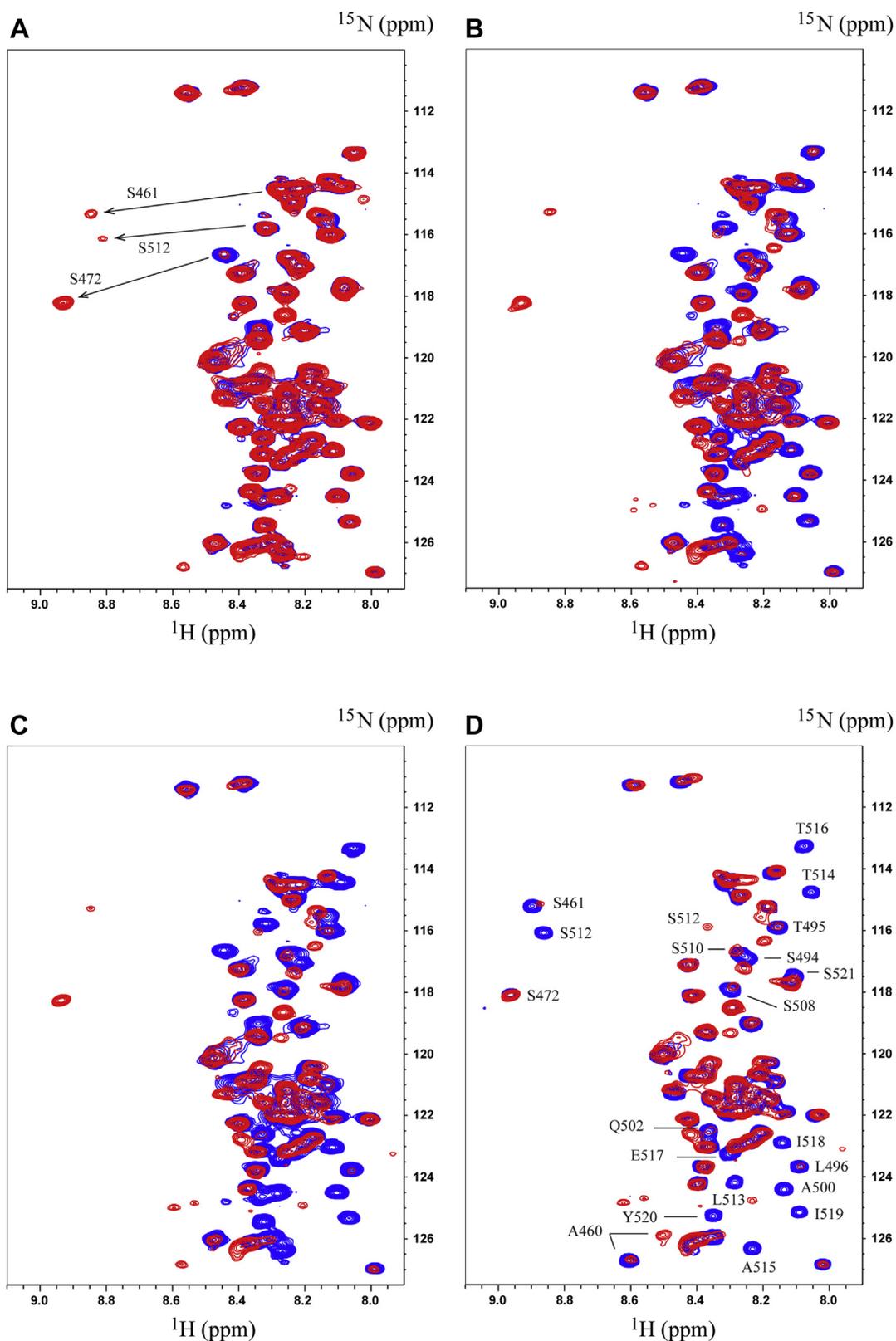


Figure 9. Longitudinal ^1H - ^{15}N 2D spectra of CEACAM1-LF peptide mixed with β -catenin after addition of PKA and comparison to the spectrum without β -catenin. A–C, the times are 48, 68, and 74 h (blue = initial spectra, red = spectra at indicated times). D, spectra overlay of fully phosphorylated CEACAM1-LF by PKA in the absence of β -catenin (blue) and in the presence of β -catenin at reaction time point of 74 h (red).

Phosphorylation of CEACAM1 by PKA and GSK3 β

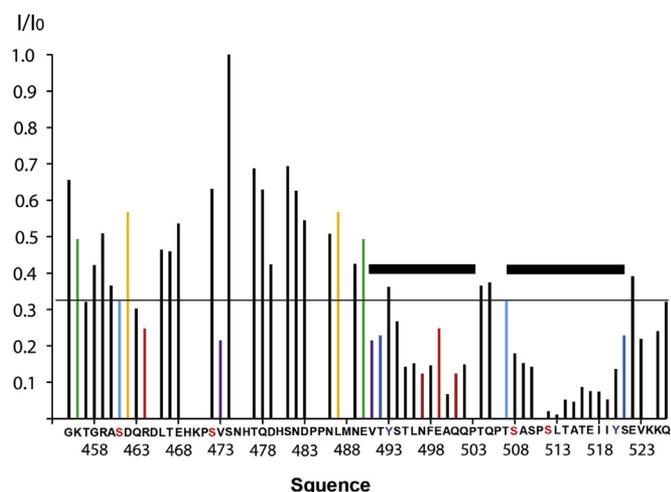


Figure 10. Peak intensity changes in the 2D NMR spectrum due to β -catenin binding to CEACAM1-LF peptide phosphorylated by PKA. Normalized peak intensities (I) taken from Figure 9C were divided by reference peak intensities in the absence of β -catenin (I_0) and plotted versus the peptide sequence. Peaks of residues with ambiguous assignments or proline are not presented. Values from pairs of overlapping residues in the 2D NMR spectra are color coded to indicate that there may be an error associated with their intensity calculation. The horizontal line is the averaged reduced peak intensity from all counted peaks. Key tyrosines (blue letters) and phosphorylated serines (red letters) are indicated along with bars showing regions most affected by interaction with β -catenin in which the peak intensity of the residues in these two regions was significantly reduced.

ITIMs that can inhibit positive ITAM signaling pathways by recruiting tandem SH2-possessing tyrosine phosphatases. Although the phosphorylation of CEACAM1-LF ITIMs by Src kinases and subsequent recruitment of dual SH2 phosphatases have been well studied, the regulation of these effectors by other sequences within the cytoplasmic domain has been less studied and may be expected based on other signaling receptors possessing ITAMs such as FcR γ (48) or ITIMs such as PECAM-1 (49). In the case of PECAM-1, that like CEACAM1-LF has two ITIMs in its cytoplasmic domain, phosphorylation of Tyr686 in its distal ITIM occurs first by the Src kinase Lyn, followed by phosphorylation of Tyr663 in its proximal ITIM by Csk or Btk. In addition, phosphorylation of Ser702 in PECAM-1 strengthens the association of residues 682 to 702 with phospholipids and their conversion from an unstructured to an alpha-helical conformation (50). In the case of rodent CEACAM1-LF, a role for the phosphorylation of Ser503 between Tyr488 and Tyr515, which define the dual ITIMs, has been based on mutational analysis of Ser503Ala that affected bile acid transport (29), insulin clearance (33), and the development of hepatosteatosis (19, 32). In those studies it was concluded that the Ser503Ala mutation negatively regulated the function of Tyr488 in the proximal ITIM and phenotypically behaved as dominant negative mutation with a similar phenotype as the *Ceacam1* KO mouse (33). Whether the same is true in humans has never been examined.

In this study, we first compared the rodent and human sequences and observed that the consensus sequence for the reported PKC phosphorylation of rodent Ser503 (29) was absent in the human equivalent residue Ser508 (Fig. 1). Instead,

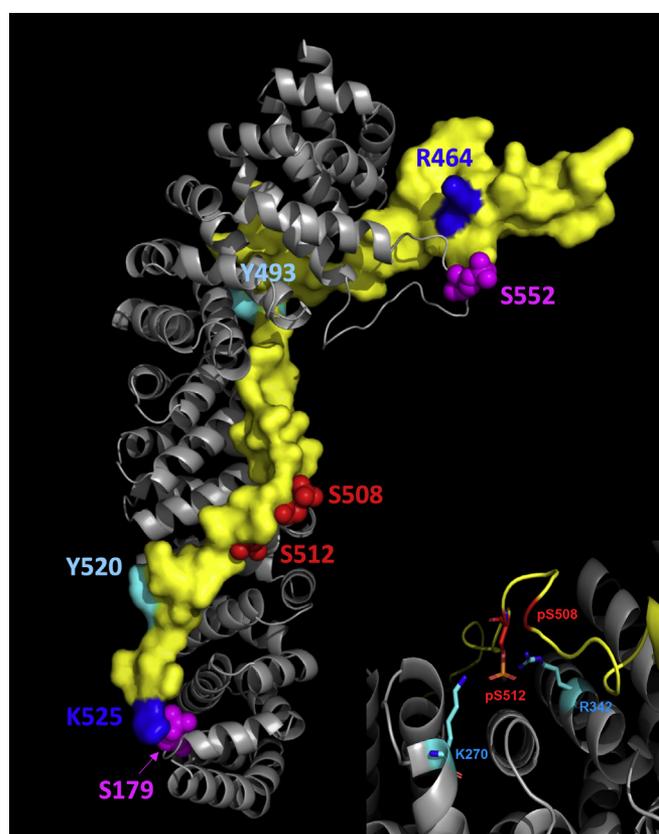


Figure 11. Energy minimized binding of PKA + GSK3 β phosphorylated CEACAM1-LF cytoplasmic domain to PKA phosphorylated β -catenin. Ribbon model of β -catenin (gray) showing locations of S552 and S179 residues phosphorylated by PKA in magenta, together with energy minimized docked structure of CEACAM1-LF cytoplasmic domain (yellow surface) with locations of S508 and S512 phosphorylated by PKA plus GSK3 β . Also shown are locations of Y493 and Y520 (cyan) and likely anchor residues R464 and K525 of CEACAM1-LF cytoplasmic domain. Inset, details of the interaction of pSer512 of CEACAM1-LF with basic residues K270 and R342 of β -catenin (distances are 2.9 and 2.8 Å, respectively).

we identified a consensus site for GSK3 β at Ser508 that was conserved in rodents and humans. Since this site required prior phosphorylation of Ser512, four amino acids downstream from Ser508, we first searched for a kinase that could perform this phosphorylation by expressing 15 N-labeled and unlabeled versions of the entire cytoplasmic domain of human CEACAM1-LF. While PKC gave no detectable phosphorylation of the peptide, PKA was able to phosphorylate three residues, one of which was Ser512. Although the unphosphorylated peptide was not a substrate for GSK3 β , when the PKA phosphorylated peptide was treated with GSK3 β , Ser508 was phosphorylated, confirming the function of the consensus site as a *bona fide* GSK3 β site. It is important to note that phosphorylation of downstream Ser512 was not observed upon treatment of the peptide with several casein kinase isoforms or AKT1 (Table S1), thus distinguishing this site from the GSK3 β priming sites in β -catenin (51) or glycogen synthase (52).

When we treated the peptide with Src, phosphorylation of Tyr493 in the proximal ITIM and Tyr520 in the distal ITIM occurred with kinetics that favored Tyr520 (Table 1). This result is similar to PECAM-1 in which phosphorylation of the distal ITIM was favored over the proximal ITIM, but not to the

Phosphorylation of CEACAM1 by PKA and GSK3 β

same extreme in which other kinases preferred the proximal ITIM in PECAM1. There was no further effect of pSer508/512 on the activity of Src toward either ITIM. Since SHP2 is the preferred phosphotyrosine phosphatase for CEACAM1 in hepatocytes (27), SHP2 was used to study the dephosphorylation kinetics. The dephosphorylation kinetics by SHP2 on the double ITIM phosphorylated peptide were similar but favored dephosphorylation of the distal ITIM when pSer508/512 was present. Since SHP2 may bind both ITIMs with its dual SH2 domains, it can be questioned if its additional role, namely dephosphorylation of the phosphorylated ITIMs, is biologically relevant. Since this is an *in vitro* study, it only answers the question of what can potentially occur in the absence of other effectors present in a biological system. However, we can conclude that Ser508/512 does play a role in the phosphorylation and dephosphorylation kinetics of the ITIMs, likely conserving a functional role for Ser508 in humans similar to Ser503 in rodents.

β -Catenin has been reported to associate with CEACAM1-LF by us (11, 12) and others (14, 53, 54), and as we now show, their association requires phosphorylation of both CEACAM1-LF and β -catenin by PKA. The results suggest that concerted regulation of CEACAM1-LF and β -catenin may occur. Three major conclusions can be made from this study. First, there was no association until the two proteins were mixed and treated with PKA, demonstrating that PKA phosphorylation of β -catenin was also required for their interaction. Since PKA phosphorylated CEACAM1-LF peptide had no measurable binding to unphosphorylated β -catenin, a concerted interaction is indicated. Second, the addition of β -catenin to Src phosphorylated or SHP2 dephosphorylated CEACAM1-LF had no further effect on their kinetics, indicating that the tyrosines were not affected by the binding of β -catenin. Third, the residues in CEACAM1-LF that interacted with β -catenin were located between the two ITIMs. Thus, the region of interaction is different than the region identified by us in the yeast two hybrid and mutational analysis (11). While the significance of the identification of a new β -catenin interaction site in CEACAM1-LF is unclear at this time, it is possible that other kinases and/or interactors present in the yeast or Jurkat cells study led to identification of a proximal binding site, *versus* the *in vitro* binding studies performed here that relied on PKA and GSK3 β phosphorylation only. The updated interaction shown in Figure 12 summarizes the results of the current study. The newly identified β -catenin binding site of CEACAM1-LF includes and lies between the two ITIMs, and except for the extreme C-terminus, is rich in acidic and

devoid of basic residues. Given the basic nature of the β -catenin-binding groove (55), this unstructured region is well suited for binding to β -catenin. It should be noted that the previously identified proximal region (Fig. 1) contains a number of basic residues that make that region a poorer candidate for binding β -catenin.

Conclusion

We have identified three PKA phosphorylation sites in the cytoplasmic domain of human CEACAM-LF, of which Ser512 was shown to be a priming site for phosphorylation of upstream Ser508 by GSK3 β . This consensus sequence is highly conserved between rodent and human CEACAM1-LF, suggesting a possible conserved function from rodent to human. Kinetics studies demonstrated an influence of pSer508/512 phosphorylation on the kinase activity of Src on the ITIM tyrosines of CEACAM1-LF and the tyrosine phosphatase activity of SHP2. NMR studies revealed a binding site for β -catenin between the ITIMs that depended on the phosphorylation of both β -catenin and CEACAM1-LF on PKA.

Experimental procedures

Peptide synthesis

A 37 amino acid long synthetic peptide of CEACAM1-LF proximal region (E490-Q526) was synthesized by the City of Hope Peptide Synthesis Core Facility with N-terminal capped with an acetyl group. The Ser508 and Ser512 were ¹⁵N-labeled by using ¹⁵N-Fmoc-Ser (Cambridge Isotope Labs). The synthesized peptide was purified by HPLC, and the molecular mass was confirmed by mass spectrometry.

Expression of CEACAM1-LF peptide and mCherry- β -catenin

The 74 amino acid (H453-Q526) cytoplasmic domain of CEACAM1-LF was expressed using *E. coli* codon-optimized overlapping oligo DNA primers and cloned into pET32b. The pET32b was modified with insertion of SMT3 to create an Ulp1 cleavage sites. The recombinant proteins were expressed in *E. coli* or *E. coli* C41(DE3) as a thioredoxin (Trx)-SMT3 fusion protein with an N-terminal His-tag. The expressed protein was first purified using a Ni-NTA resin, and further purified by cleaving the peptide off the N-terminal Trx-SMT3 portion by incubation with hexa-histidine tagged Ulp1. Expression of the recombinant protein and cleaved peptides were confirmed by SDS-PAGE. The protein was concentrated using Centricon YM3 centrifugal filters (SigmaAldrich) and purified by FPLC on a Superdex 75G column (GE Healthcare Life Science) using 50 mM Tris buffer with 150 mM NaCl and



Figure 12. Phosphorylation sites and β -catenin binding in CEACAM1-LF cytoplasmic domain. Phosphorylation sites identified for PKA, GSK3 β , and Src are shown with arrows. Acidic residues (D and E) plus the pSer residues identified are shown in red, basic residues R and K in blue and H in green. The β -catenin-binding domain identified by NMR boxed in blue.

Phosphorylation of CEACAM1 by PKA and GSK3 β

10% glycerol at pH 8. The purified protein was buffer exchanged to pH 6.6, 30 mM phosphate buffer and the sequence confirmed by mass spectrometry. The ^{15}N -labeled CEACAM1-LF was expressed using M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ as sole nitrogen source, and ^{15}N , ^{13}C -labeled CEACAM1-LF was expressed using M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ and D-Glucose ($\text{U-}^{13}\text{C}_6$).

In order to increase the solubility of the armadillo repeats of β -catenin (residues 138–666) at pH 6.6, it was expressed as fusion protein with mCherry1, with the acidic linker GDEV-DEDEG followed by a His₆ tag LEHHHHHH. The DNA encoding this construct, flanked by NcoI and XhoI cloning sites, was assembled as a synthetic gene from GenArt (Life Technologies/Thermo Fisher) and cloned into a pET28b vector. The plasmid insert was validated by DNA sequencing, and the fusion protein was expressed in *E. coli* (*E. coli* express B21(DE3) cells). The fusion protein was purified by Ni-NTA affinity and size-exclusion chromatography, and its purity and molecular mass validated by SDS gel electrophoresis and mass spectrometry.

NMR sample preparation

For the backbone assignment of CEACAM1-LF, the ^{15}N -labeled and ^{15}N -, ^{13}C -labeled samples were prepared in 50 mM NaAc-d₃ buffer (5% D₂O) with pH 5.5. For the kinase assay study, the CEACAM1-LF peptide was prepared in 5% D₂O, pH 6.6, 30 mM phosphate buffer containing 100 mM D-mannitol. The same buffer was also used for the NMR sample preparation of mCherry- β -catenin. The concentration of stock peptide was calibrated using TSP-d₄ as internal reference using ^1H -NMR, and the mCherry- β -catenin concentration was determined 280 nm absorbance.

NMR experiments and data analysis

All experiments were carried out at 25 °C on a 700 MHz Bruker Ascend with a TXI-triple resonance cryoprobe. For the backbone assignment of CEACAM1-LF cytoplasmic domain peptide, 2D ^1H - ^{15}N -HSQC, ^1H - ^{13}C -HSQC, 3D ^{15}N -TOCSY-HSQC, 3D ^{15}N -NOESY-HSQC, HCC(CO)NH₂(H)CC(CO)NH₂, HNCO, HNCACO, HNCACB, HNCOCACB data were acquired. The sample concentration used for backbone assignment experiment is 0.4 mM for ^{15}N , ^{13}C -CEACAM1-LF, and 0.9 mM for ^{15}N -CEACAM1-LF. The data were processed and analyzed using NMRPipe (56), NMRView (57), NMRFAM-SPARKY (58), and Bruker Topspin. The CEACAM1-LF amide assignment at pH 6.6 was obtained from that at pH 5.5 through monitoring ^1H - ^{15}N -HSQC spectra at different pH titrated from pH 5.5 to pH 6.6. The longitudinal ^1H - ^{15}N spectra was acquired using sofast HMQC (59) to monitor the enzyme kinetics.

Kinase assays

The kinases and phosphatases used in this study are listed in Table S2. For kinase assays monitored by 2D ^1H - ^{15}N sofast HMQC, the typical ^{15}N -CEACAM1-LF concentration is

around 20 μM . The 30 mM phosphate buffer contains 5% D₂O, 100 mM D-mannitol, 0.5 mM ATP, 5 mM MgCl₂, 1 mM TCEP, and 16 μl EDTA-free protease inhibitor cocktail solution (one mini Roche tablet dissolved in 1.5 ml 30 mM phosphate buffer). To 500 μl ^{15}N -CEACAM1-LF peptide, different kinases were added with calculated amount based on the enzyme specific activity so that the reaction could be finished in about 6 h. After ^{15}N -CEACAM1-LF peptide was phosphorylated, the kinase was denatured by heating the sample at 60 °C for 30 min. The denatured kinase was separated from phosphorylated ^{15}N -CEACAM1-LF using a 30 kDa cutoff centrifugal filter, and the flow-through phosphorylated ^{15}N -CEACAM1-LF peptide was then exchanged to 30 mM phosphate buffer for further phosphorylation with other kinases or treatment with SHP2 phosphatase. For the SHP2 phosphatase assay, the 30 mM phosphate buffer contains 5% D₂O, 100 mM D-mannitol, 1 mM TCEP, 0.5 mM EDTA, and 16 μl EDTA-free protease inhibitor cocktail solution. Some reaction monitoring was carried out in 3 mm NMR tube with 180 μl sample volume instead of 500 μl . All the kinase and phosphatase assays were carried out at 25 °C. The enzyme apparent reaction rate k_{app} is defined by $k_{\text{app}} = k_{\text{cat}} \times [E]_{t=0} / K_{\text{M}}$, where $[E]_{t=0}$ is initial enzyme concentration, k_{cat} is the turnover numbers of substrate, and K_{M} is Michaelis–Menten constants of kinase (60). The cross peak intensity versus time was fitted to equation of $[S]_t = ([S]_{t=0} - C)\exp(-k_{\text{app}} \times t) + C$ to derive the k_{app} , where the t stands for time, $[S]_t$ stands for peptide concentration at time point t . C is a constant.

The unlabeled CEACAM1-LF peptide was phosphorylated for mass spectroscopy as follows: The peptide (600 μl of 15 μM stock) was prepared in 5% D₂O, pH 6.6, 30 mM phosphate buffer together with 1 mM ATP, 6 mM MgCl₂, 1 mM TCEP and 24 μl of EDTA-free protease inhibitor cocktail solution (Roche). The one-dimensional ^1H NMR spectrum was acquired right after addition of 5 μl PKA (2.2 μg ; Sigma Aldrich, catalog 14-440). The reaction was monitored over time using 1D ^1H NMR spectrum by observing new peaks that appeared at low field of amide region. The amide protons of Ser or Thr move downfield by 0.5 to 1.5 ppm when they are phosphorylated. The reaction was stopped after 5 days by heating the sample at 60 °C for 30 min. A 100 μl aliquot was taken out from NMR tube for mass spectrometry study. GSK3 β kinase (4 μl , 1.2 μg ; Thermo Fisher, catalog PV3365) plus ATP (3 μl of 46 mM) and MgCl₂ (3 μl of 250 mM) were added to the remaining sample. The reaction was carried out at room temperature and monitored by 1D ^1H NMR for changes at amide region. After 2 days reaction, a 100 μl aliquot was taken out from NMR tube for sequence analysis by mass spectrometry.

MALDI-MS and LC-ESI-MS/MS analysis of phosphorylated CEACAM1-LF peptide

Kinase-treated peptides (160 ng, see Kinase assays above) were analyzed by MALDI-MS on a Shimadzu MALDI-TOF MS 8020 or digested in-solution by filter-aided sample preparation using 5 kDa molecular weight cutoff filters (Amicon, Millipore

Sigma) using rLys-C (Promega, 1:50 enzyme/protein ratio). The resulting peptides were desalted and concentrated with Oasis HLB 1 cc 10-mg cartridges (Waters), eluted with 70% acetonitrile (ACN) in 0.1% TFA, freeze-dried completely, and resuspended in 0.1% formic acid. Samples were analyzed by high-resolution nano-electrospray LC-MS and MS/MS using an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo) equipped with an EasyNano LC in positive-ion mode. Buffers: solution A was 0.1% formic acid in water, and solution B was 0.1% formic acid in acetonitrile (98% acetonitrile). Five microliters of each sample was loaded onto an Acclaim PepMap 100 C18 LC trapping column (3 μ m, 75 μ m \times 2 cm, 100 \AA pores, Thermo Fisher Scientific) at a flow rate of 5 μ l/min with buffer A and separated by an analytical EASY-Spray column, 25 cm \times 75 μ m ID, PepMap RSLC C18 2 μ m (Thermo) at a flow rate of 500 nl/min. Samples were eluted using a gradient starting with a linear increase from 3 to 25% of buffer B (acetonitrile containing 0.1% formic acid) acetonitrile over 40 min, followed by an increase from 25 to 90% acetonitrile and column temperature of 45 $^{\circ}$ C, which was maintained until 45 min. Acquisition analysis duration was 45 min using neutral-loss-triggered fragmentation and multistage activation approach for the analysis of phosphopeptides (61). The following MS settings were used: an acquisition time of 45 min, a positive-ion spray voltage of 2300 V, an ion-transfer-tube temperature of 275 $^{\circ}$ C. MS¹ was acquired in the orbitrap at 120,000 resolution in the m/z range 400 to 1600, the AGC target was 400,000, the S-lens RF level was set to 60%. MS² spectra were acquired in the Orbitrap at 30,000 resolution with quadrupole-isolation mode and a 0.7 Da isolation window. Each precursor was fragmented using CID, ETD, and HCD with a normalized collision energy of 35%. One microscan was acquired with an activation Q of 0.25. The data were analyzed using PEAKS X⁺ with FDR of 2.0% for peptide-spectrum matches and the parent-mass-error tolerance set to 10.0 ppm and fragment-mass-error tolerance set to 0.05 Da. Phosphorylation sites in MS² spectra were assigned using Ascores from 10 to 20.

In silico modeling of CEACAM1-LF/ β -catenin interaction complex

A model of the CEACAM1-LF/ β -catenin complex was predicted using homology modeling, based on available crystal structures of β -catenin bound protein complexes as templates: (a) phosphorylated human APC/ β -catenin (PDB ID: 1V18) (62), (b) phosphorylated human E-cadherin/ β -catenin (PDB ID: 1I7W) (47), (c) human TCF4/ β -catenin (PDB ID: 1JPW) (63) and (d) XTFC3/human β -catenin, TCF3 from *Xenopus laevis* (PDB ID: 1G3J) (64). The sequence alignment of the four templates was achieved by an initial structural alignment in PyMOL (65), followed by structure guided sequence alignment in Chimera (66). Next, the sequence of the cytoplasmic domain of CEACAM1-LF was aligned to the four templates using Chimera. To emulate phosphorylation, four residues on CEACAM1-LF sequence were mutated to Asp or Glu (Y493E, S508D, S512D, Y521E) to simulate phosphorylation of Ser or

Tyr residues. The Asp/Glu mutations were performed only for the purpose of the sequence alignment. The homology modeling was performed using Modeller (67), employing the “loopmodel” module. To model the interactions between pS552 and pS179 of β -catenin with R464 and K525 of CEACAM1-LF, harmonic distance restraints were enacted between the C α atoms of these residues, with an equilibrium distance of 8.5 \AA . The final model was selected based on the lowest molpdf score. In the final model, S179, S552 of β -catenin and S508, S512 of CEACAM1-LF were changed to their phosphorylated forms using Maestro (68) (Schrodinger). The 5 \AA environments of these four residues were optimized through side-chain reassignment using Prime (69), followed by minimization of the whole complex.

Data availability

The raw data for these studies are available from the senior author at jshively@coh.org. The mass spectrometry raw files have been deposited with PRIDE as follows: Project Name: Human CEACAM1-LF regulates lipid storage in hepatocytes via CD36 and insulin receptor expression. Project accession: PXD021463. Project DOI: [10.6019/PXD021463](https://doi.org/10.6019/PXD021463).

Supporting information—This article contains supporting information.

Acknowledgments—This research was supported by NCI grant P30CA033572. We thank Dr Yuelong Ma of the Synthetic and biopolymer chemistry core at City of Hope for peptide synthesis.

Author contributions—J. E. S. conceptualization; W. H., K. B., S. B., T. H., A. T., P. W., and M. K. data curation; W. H., S. B., M. K., and J. E. S. formal analysis; W. H., K. B., S. B., T. H., A. T., P. W., and M. K. investigation; W. H. methodology; J. E. S. resources; M. K. and J. E. S. supervision; W. H. writing—original draft; S. B., M. K., and J. E. S. writing—review and editing.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CEACAM1, carcinoembryonic antigen-related cell adhesion molecule-1; GSK3 β , glycogen synthase kinase 3 β ; ITAM, immunoreceptor tyrosine activation motif; ITIM, immunoreceptor tyrosine inhibitory motif; NMR, nuclear magnetic resonance; PKA, protein kinase A.

References

- Gray-Owen, S. D., and Blumberg, R. S. (2006) CEACAM1: Contact-dependent control of immunity. *Nat. Rev. Immunol.* **6**, 433–446
- Calinescu, A., Turcu, G., Nedelcu, R. I., Brinzea, A., Hodoroaga, A., Antohe, M., Diaconu, C., Bleotu, C., Pirici, D., Jilaveanu, L. B., Ion, D. A., and Badarau, I. A. (2018) On the dual role of carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) in human malignancies. *J. Immunol. Res.* **2018**, 7169081
- Gaur, S., Shively, J. E., Yen, Y., and Gaur, R. K. (2008) Altered splicing of CEACAM1 in breast cancer: Identification of regulatory sequences that control splicing of CEACAM1 into long or short cytoplasmic domain isoforms. *Mol. Cancer* **7**, 46

Phosphorylation of CEACAM1 by PKA and GSK3 β

- Dery, K. J., Kujawski, M., Grunert, D., Wu, X., Ngyuen, T., Cheung, C., Yim, J. H., and Shively, J. E. (2014) IRF-1 regulates alternative mRNA splicing of carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) in breast epithelial cells generating an immunoreceptor tyrosine-based inhibition motif (ITIM) containing isoform. *Mol. Cancer* **13**, 64
- Brummer, J., Neumaier, M., Gopfert, C., and Wagener, C. (1995) Association of pp60c-src with biliary glycoprotein (CD66a), an adhesion molecule of the carcinoembryonic antigen family downregulated in colorectal carcinomas. *Oncogene* **11**, 1649–1655
- Beauchemin, N., Kunath, T., Robitaille, J., Chow, B., Turbide, C., Daniels, E., and Veillette, A. (1997) Association of biliary glycoprotein with protein tyrosine phosphatase SHP-1 in malignant colon epithelial cells. *Oncogene* **14**, 783–790
- Nagaishi, T., Pao, L., Lin, S. H., Iijima, H., Kaser, A., Qiao, S. W., Chen, Z., Glickman, J., Najjar, S. M., Nakajima, A., Neel, B. G., and Blumberg, R. S. (2006) SHP1 phosphatase-dependent T cell inhibition by CEACAM1 adhesion molecule isoforms. *Immunity* **25**, 769–781
- Huber, M., Izzi, L., Grondin, P., Houde, C., Kunath, T., Veillette, A., and Beauchemin, N. (1999) The carboxyl-terminal region of biliary glycoprotein controls its tyrosine phosphorylation and association with protein-tyrosine phosphatases SHP-1 and SHP-2 in epithelial cells. *J. Biol. Chem.* **274**, 335–344
- Billadeau, D. D., and Leibson, P. J. (2002) ITAMs versus ITIMs: Striking a balance during cell regulation. *J. Clin. Invest.* **109**, 161–168
- Najjar, S. M., and Perdomo, G. (2019) Hepatic insulin clearance: Mechanism and physiology. *Physiology (Bethesda)* **34**, 198–215
- Jin, L., Li, Y., Chen, C. J., Sherman, M. A., Le, K., and Shively, J. E. (2008) Direct interaction of tumor suppressor CEACAM1 with beta catenin: Identification of key residues in the long cytoplasmic domain. *Exp. Biol. Med. (Maywood)* **233**, 849–859
- Li, Y., and Shively, J. E. (2013) CEACAM1 regulates Fas-mediated apoptosis in Jurkat T-cells via its interaction with beta-catenin. *Exp. Cell Res.* **319**, 1061–1072
- Leung, N., Turbide, C., Balachandra, B., Marcus, V., and Beauchemin, N. (2008) Intestinal tumor progression is promoted by decreased apoptosis and dysregulated Wnt signaling in Ceacam1^{-/-} mice. *Oncogene* **27**, 4943–4953
- Wegwitz, F., Lenfert, E., Gerstel, D., von Ehrenstein, L., Einhoff, J., Schmidt, G., Logsdon, M., Brandner, J., Tiegs, G., Beauchemin, N., Wagener, C., Deppert, W., and Horst, A. K. (2016) CEACAM1 controls the EMT switch in murine mammary carcinoma *in vitro* and *in vivo*. *Oncotarget* **7**, 63730–63746
- Bienz, M. (2005) beta-Catenin: A pivot between cell adhesion and Wnt signalling. *Curr. Biol.* **15**, R64–67
- Valenta, T., Hausmann, G., and Basler, K. (2012) The many faces and functions of beta-catenin. *EMBO J.* **31**, 2714–2736
- Leung, N., Turbide, C., Olson, M., Marcus, V., Jothy, S., and Beauchemin, N. (2006) Deletion of the carcinoembryonic antigen-related cell adhesion molecule 1 (Ceacam1) gene contributes to colon tumor progression in a murine model of carcinogenesis. *Oncogene* **25**, 5527–5536
- Russo, L., Muturi, H. T., Ghadieh, H. E., Ghanem, S. S., Bowman, T. A., Noh, H. L., Dagdeviren, S., Dogbey, G. Y., Kim, J. K., Heinrich, G., and Najjar, S. M. (2017) Liver-specific reconstitution of CEACAM1 reverses the metabolic abnormalities caused by its global deletion in male mice. *Diabetologia* **60**, 2463–2474
- Dai, T., Abou-Rjaily, G. A., Al-Share, Q. Y., Yang, Y., Fernstrom, M. A., Deangelis, A. M., Lee, A. D., Sweetman, L., Amato, A., Pasquali, M., Lopaschuk, G. D., Erickson, S. K., and Najjar, S. M. (2004) Interaction between altered insulin and lipid metabolism in CEACAM1-inactive transgenic mice. *J. Biol. Chem.* **279**, 45155–45161
- Chen, C. J., Kirshner, J., Sherman, M. A., Hu, W., Nguyen, T., and Shively, J. E. (2007) Mutation analysis of the short cytoplasmic domain of the cell-cell adhesion molecule CEACAM1 identifies residues that orchestrate actin binding and lumen formation. *J. Biol. Chem.* **282**, 5749–5760
- Kirshner, J., Schumann, D., and Shively, J. E. (2003) CEACAM1, a cell-cell adhesion molecule, directly associates with annexin II in a three-dimensional model of mammary morphogenesis. *J. Biol. Chem.* **278**, 50338–50345
- Hu, W., Bhattacharya, S., Hong, T., Wong, P., Li, L., Vaidehi, N., Kalkum, M., and Shively, J. E. (2021) Structural characterization of a dimeric complex between the short cytoplasmic domain of CEACAM1 and the pseudo tetramer of S100A10-Annexin A2 using NMR and molecular dynamics. *Biochim. Biophys. Acta Biomembr.* **1863**, 183451
- Kirshner, J., Chen, C. J., Liu, P., Huang, J., and Shively, J. E. (2003) CEACAM1-4S, a cell-cell adhesion molecule, mediates apoptosis and reverts mammary carcinoma cells to a normal morphogenic phenotype in a 3D culture. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 521–526
- Yokoyama, S., Chen, C. J., Nguyen, T., and Shively, J. E. (2007) Role of CEACAM1 isoforms in an *in vivo* model of mammary morphogenesis: Mutational analysis of the cytoplasmic domain of CEACAM1-4S reveals key residues involved in lumen formation. *Oncogene* **26**, 7637–7646
- Hunter, I., Sawa, H., Edlund, M., and Obrink, B. (1996) Evidence for regulated dimerization of cell-cell adhesion molecule (C-CAM) in epithelial cells. *Biochem. J.* **320**, 847–853
- Ocklind, C., Forsum, U., and Obrink, B. (1983) Cell surface localization and tissue distribution of a hepatocyte cell-cell adhesion glycoprotein (cell-CAM 105). *J. Cell Biol.* **96**, 1168–1171
- Muller, M. M., Klaile, E., Vorontsova, O., Singer, B. B., and Obrink, B. (2009) Homophilic adhesion and CEACAM1-S regulate dimerization of CEACAM1-L and recruitment of SHP-2 and c-Src. *J. Cell Biol.* **187**, 569–581
- Kammerer, R., and Zimmermann, W. (2010) Coevolution of activating and inhibitory receptors within mammalian carcinoembryonic antigen families. *BMC Biol.* **8**, 12
- Sippel, C. J., Fallon, R. J., and Perlmutter, D. H. (1994) Bile acid efflux mediated by the rat liver canalicular bile acid transport/ecto-ATPase protein requires serine 503 phosphorylation and is regulated by tyrosine 488 phosphorylation. *J. Biol. Chem.* **269**, 19539–19545
- Ghosh, S., Kaw, M., Patel, P. R., Ledford, K. J., Bowman, T. A., McInerney, M. F., Erickson, S. K., Bourey, R. E., and Najjar, S. M. (2010) Mice with null mutation of Ceacam I develop nonalcoholic steatohepatitis. *Hepat. Med.* **2010**, 69–78
- Park, S. Y., Cho, Y. R., Kim, H. J., Hong, E. G., Higashimori, T., Lee, S. J., Goldberg, I. J., Shulman, G. I., Najjar, S. M., and Kim, J. K. (2006) Mechanism of glucose intolerance in mice with dominant negative mutation of CEACAM1. *Am. J. Physiol. Endocrinol. Metab.* **291**, E517–E524
- Lee, S. J., Heinrich, G., Fedorova, L., Al-Share, Q. Y., Ledford, K. J., Fernstrom, M. A., McInerney, M. F., Erickson, S. K., Gatto-Weis, C., and Najjar, S. M. (2008) Development of nonalcoholic steatohepatitis in insulin-resistant liver-specific S503A carcinoembryonic antigen-related cell adhesion molecule 1 mutant mice. *Gastroenterology* **135**, 2084–2095
- Poy, M. N., Yang, Y., Rezaei, K., Fernstrom, M. A., Lee, A. D., Kido, Y., Erickson, S. K., and Najjar, S. M. (2002) CEACAM1 regulates insulin clearance in liver. *Nat. Genet.* **30**, 270–276
- DeAngelis, A. M., Heinrich, G., Dai, T., Bowman, T. A., Patel, P. R., Lee, S. J., Hong, E. G., Jung, D. Y., Assmann, A., Kulkarni, R. N., Kim, J. K., and Najjar, S. M. (2008) Carcinoembryonic antigen-related cell adhesion molecule 1: A link between insulin and lipid metabolism. *Diabetes* **57**, 2296–2303
- Frame, S., Cohen, P., and Biondi, R. M. (2001) A common phosphate binding site explains the unique substrate specificity of GSK3 and its inactivation by phosphorylation. *Mol. Cell* **7**, 1321–1327
- Kim, K. H., Song, M. J., Yoo, E. J., Choe, S. S., Park, S. D., and Kim, J. B. (2004) Regulatory role of glycogen synthase kinase 3 for transcriptional activity of ADD1/SREBP1c. *J. Biol. Chem.* **279**, 51999–52006
- Svenberg, T., Hammarstrom, S., and Hedin, A. (1979) Purification and properties of biliary glycoprotein I (BGP I). Immunochemical relationship to carcinoembryonic antigen. *Mol. Immunol.* **16**, 245–252
- Hinoda, Y., Neumaier, M., Hefta, S. A., Drzeniek, Z., Wagener, C., Shively, L., Hefta, L. J., Shively, J. E., and Paxton, R. J. (1988) Molecular cloning of a cDNA coding biliary glycoprotein I: Primary structure of a glycoprotein immunologically crossreactive with carcinoembryonic antigen. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6959–6963

39. Dawson, P. A., Lan, T., and Rao, A. (2009) Bile acid transporters. *J. Lipid Res.* **50**, 2340–2357
40. Shanks, N., Greek, R., and Greek, J. (2009) Are animal models predictive for humans? *Philos. Ethics Humanit. Med.* **4**, 2
41. Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241
42. Theillet, F. X., Smet-Nocca, C., Liokatis, S., Thongwichian, R., Kosten, J., Yoon, M. K., Kriwacki, R. W., Landrieu, I., Lippens, G., and Selenko, P. (2012) Cell signaling, post-translational protein modifications and NMR spectroscopy. *J. Biomol. NMR* **54**, 217–236
43. Du, J. T., Li, Y. M., Wei, W., Wu, G. S., Zhao, Y. F., Kanazawa, K., Nemoto, T., and Nakanishi, H. (2005) Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide. *J. Am. Chem. Soc.* **127**, 16350–16351
44. Fang, D., Hawke, D., Zheng, Y., Xia, Y., Meisenhelder, J., Nika, H., Mills, G. B., Kobayashi, R., Hunter, T., and Lu, Z. (2007) Phosphorylation of beta-catenin by AKT promotes beta-catenin transcriptional activity. *J. Biol. Chem.* **282**, 11221–11229
45. Hino, S., Tanji, C., Nakayama, K. I., and Kikuchi, A. (2005) Phosphorylation of beta-catenin by cyclic AMP-dependent protein kinase stabilizes beta-catenin through inhibition of its ubiquitination. *Mol. Cell. Biol.* **25**, 9063–9072
46. Taurin, S., Sandbo, N., Qin, Y., Browning, D., and Dulin, N. O. (2006) Phosphorylation of beta-catenin by cyclic AMP-dependent protein kinase. *J. Biol. Chem.* **281**, 9971–9976
47. Huber, A. H., and Weis, W. I. (2001) The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* **105**, 391–402
48. Getahun, A., and Cambier, J. C. (2015) Of ITIMs, ITAMs, and ITAMis: Revisiting immunoglobulin Fc receptor signaling. *Immunol. Rev.* **268**, 66–73
49. Tourdot, B. E., Brenner, M. K., Keough, K. C., Holyst, T., Newman, P. J., and Newman, D. K. (2013) Immunoreceptor tyrosine-based inhibitory motif (ITIM)-mediated inhibitory signaling is regulated by sequential phosphorylation mediated by distinct nonreceptor tyrosine kinases: A case study involving PECAM-1. *Biochemistry* **52**, 2597–2608
50. Paddock, C., Lytle, B. L., Peterson, F. C., Holyst, T., Newman, P. J., Volkman, B. F., and Newman, D. K. (2011) Residues within a lipid-associated segment of the PECAM-1 cytoplasmic domain are susceptible to inducible, sequential phosphorylation. *Blood* **117**, 6012–6023
51. Wu, D., and Pan, W. (2010) GSK3: A multifaceted kinase in Wnt signaling. *Trends Biochem. Sci.* **35**, 161–168
52. Beurel, E., Grieco, S. F., and Jope, R. S. (2015) Glycogen synthase kinase-3 (GSK3): Regulation, actions, and diseases. *Pharmacol. Ther.* **148**, 114–131
53. Ghavampour, S., Kleefeldt, F., Bommel, H., Volland, J., Paus, A., Horst, A., Pfeiffer, V., Hubner, S., Wagner, N., Rueckschloss, U., and Ergun, S. (2018) Endothelial barrier function is differentially regulated by CEACAM1-mediated signaling. *FASEB J.* **32**, 5612–5625
54. Sundberg, U., Beauchemin, N., and Obrink, B. (2004) The cytoplasmic domain of CEACAM1-L controls its lateral localization and the organization of desmosomes in polarized epithelial cells. *J. Cell Sci.* **117**, 1091–1104
55. Huber, A. H., Nelson, W. J., and Weis, W. I. (1997) Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**, 871–882
56. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293
57. Johnson, B. A., and Blevins, R. A. (1994) NMR view: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* **4**, 603–614
58. Lee, W., Tonelli, M., and Markley, J. L. (2015) NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327
59. Schanda, P., and Brutscher, B. (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. *J. Am. Chem. Soc.* **127**, 8014–8015
60. Theillet, F. X., Rose, H. M., Liokatis, S., Binolfi, A., Thongwichian, R., Stuijver, M., and Selenko, P. (2013) Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat. Protoc.* **8**, 1416–1432
61. Potel, C. M., Lemeer, S., and Heck, A. J. R. (2019) Phosphopeptide fragmentation and site localization by mass spectrometry: An update. *Anal. Chem.* **91**, 126–141
62. Ha, N. C., Tonzuka, T., Stamos, J. L., Choi, H. J., and Weis, W. I. (2004) Mechanism of phosphorylation-dependent binding of APC to beta-catenin and its role in beta-catenin degradation. *Mol. Cell* **15**, 511–521
63. Poy, F., Lepourcelet, M., Shivdasani, R. A., and Eck, M. J. (2001) Structure of a human Tcf4-beta-catenin complex. *Nat. Struct. Biol.* **8**, 1053–1057
64. Graham, T. A., Weaver, C., Mao, F., Kimelman, D., and Xu, W. (2000) Crystal structure of a beta-catenin/Tcf complex. *Cell* **103**, 885–896
65. Schrodinger, LLC. (2015) *The PyMOL Molecular Graphics System, Version 1.8*, DeLano Scientific LLC, South San Francisco, CA
66. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612
67. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
68. *Schrödinger Release 2018*. (2018). Maestro, Schrödinger, LLC, New York, NY
69. Jacobson, M. P., Friesner, R. A., Xiang, Z., and Honig, B. (2002) On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **320**, 597–608