

## Original Article

# Biochemical features of the adhesion G protein-coupled receptor CD97 related to its auto-proteolysis and HeLa cell attachment activities

Li-yun YANG<sup>1,2</sup>, Xiao-fang LIU<sup>1</sup>, Yang YANG<sup>1</sup>, Lin-lin YANG<sup>1</sup>, Kai-wen LIU<sup>2</sup>, Yu-bo TANG<sup>1</sup>, Min ZHANG<sup>1</sup>, Min-jia TAN<sup>1</sup>, Shan-mei CHENG<sup>1</sup>, Ye-chun XU<sup>1</sup>, Huai-yu YANG<sup>1</sup>, Zhi-jie LIU<sup>2</sup>, Gao-jie SONG<sup>2,\*</sup>, Wei HUANG<sup>1,2,\*</sup>

<sup>1</sup>CAS Key Laboratory of Receptor Research, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China; <sup>2</sup>iHuman Institute, Shanghai Tech University, Shanghai 201210, China

### Abstract

CD97 belongs to the adhesion GPCR family characterized by a long ECD linked to the 7TM via a GPCR proteolytic site (GPS) and plays important roles in modulating cell migration and invasion. CD97 (EGF1-5) is a splicing variant of CD97 that recognizes a specific ligand chondroitin sulfate on cell membranes and the extracellular matrix. The aim of this study was to elucidate the extracellular molecular basis of the CD97 EGF1-5 isoform in protein expression, auto-proteolysis and cell adhesion, including epidermal growth factor (EGF)-like domain, GPCR autoproteolysis-inducing (GAIN) domain, as well as GPS mutagenesis and N-glycosylation. Both wild-type (WT) CD97-ECD and its truncated, GPS mutated, PNGase F-deglycosylated, and N-glycosylation site mutated forms were expressed and purified. The auto-proteolysis of the proteins was analyzed with Western blotting and SDS-PAGE. Small angle X-ray scattering (SAXS) and molecular modeling were used to determine a structural profile of the properly expressed receptor. Potential N-glycosylation sites were identified using MS and were modulated with PNGase F digestion and glyco-site mutations. A flow cytometry-based HeLa cell attachment assay was used for all aforementioned CD97 variants to elucidate the molecular basis of CD97-HeLa interactions. A unique concentration-dependent GPS auto-proteolysis was observed in CD97 EGF1-5 isoform with the highest concentration (4 mg/mL) per sample was self-cleaved much faster than the lower concentration (0.1 mg/mL), supporting an intermolecular mechanism of auto-proteolysis that is distinct to the reported intramolecular mechanism for other CD97 isoforms. N-glycosylation affected the auto-proteolysis of CD97 EGF1-5 isoform in a similar way as the other previously reported CD97 isoforms. SAXS data for WT and deglycosylated CD97ECD revealed a spatula-like shape with GAIN and EGF domains constituting the body and handle, respectively. Structural modeling indicated a potential interaction between the GAIN and EGF5 domains accounting for the absence of expression of the GAIN domain itself, although EGF5-GAIN was expressed similarly in the wild-type protein. For HeLa cell adhesion, the GAIN-truncated forms showed dramatically reduced binding affinity. The PNGase F-deglycosylated and GPS mutated forms also exhibited reduced HeLa attachment compared with WT CD97. However, neither N-glycosylation mutagenesis nor auto-proteolysis inhibition caused by N-glycosylation mutagenesis affected CD97-HeLa cell interactions. A comparison of the HeLa binding affinities of PNGase F-digested, GPS-mutated and N-glycosylation-mutated CD97 samples revealed diverse findings, suggesting that the functions of CD97 ECD were complex, and various technologies for function validation should be utilized to avoid single-approach bias when investigating N-glycosylation and auto-proteolysis of CD97. A unique mechanism of concentration-dependent auto-proteolysis of the CD97 EGF1-5 isoform was characterized, suggesting an intermolecular mechanism that is distinct from that of other previously reported CD97 isoforms. The EGF5 and GAIN domains are likely associated with each other as CD97 expression and SAXS data revealed a potential interaction between the two domains. Finally, the GAIN and EGF domains are also important for CD97-HeLa adhesion, whereas N-glycosylation of the CD97 GAIN domain and GPS auto-proteolysis are not required for HeLa cell attachment.

**Keywords:** CD97; adhesion GPCR; auto-proteolysis; HeLa cell attachment; N-glycosylation

Acta Pharmacologica Sinica (2017) 38: 56–68; doi: 10.1038/aps.2016.89; published online 19 Sep 2016

### Introduction

G protein-coupled receptors (GPCRs) share a common architecture of a seven transmembrane domain (7TM), and they

can be subdivided into five groups: rhodopsin-like (Class A), secretin (Class B), glutamate (Class C), adhesion (Class D), and frizzled/smoothed (Class F) receptors<sup>[1-3]</sup>. The five groups differ greatly in the length and structure of the extracellular domains (ECDs) and ligand recognition patterns. Among these groups of receptors, adhesion GPCRs are characterized by a long ECD linked to the 7TM via a GPCR proteolytic site

\*To whom correspondence should be addressed.

E-mail huangwei@simmm.ac.cn (Wei HUANG);

songgj@shanghaitech.edu.cn (Gao-jie SONG)

Received 2016-04-06 Accepted 2016-06-12

(GPS)<sup>[4-6]</sup>. The ECD contains tandem adhesive domains (*eg*, epidermal growth factor (EGF)-like, thrombospondin repeats (TSR), or Leucine-rich repeats (LRR)) at the N-terminus, followed by a common GPCR autoproteolysis-inducing (GAIN) domain<sup>[7,8]</sup>. Adhesion GPCRs bind various cellular and matrix ligands through these distinct adhesive domains<sup>[4,9-13]</sup>.

CD97 belongs to the adhesion GPCR family and is normally expressed in many tissues by cells such as smooth muscle cells<sup>[14]</sup>, monocytes<sup>[15]</sup>, and immune cells<sup>[16,17]</sup>. CD97 splicing variants, *ie*, CD97(EGF1,2,5), CD97(EGF1,2,3,5), and CD97(EGF1-5), contain 3–5 EGF motifs in the extracellular region<sup>[15,18]</sup>. Each isoform or domain has a specific ligand: CD97(EGF1,2,5) binds to a complementary regulatory protein CD55/DAF (decay-accelerating factor) via EGF domains 1 and 2<sup>[11,19,20]</sup>; and CD97(EGF1-5) recognizes chondroitin sulfate (CS), a glycosaminoglycan, on cell membranes and the extracellular matrix<sup>[10,21]</sup>. An arginine-glycine-aspartate (RGD) motif in the GAIN domain of all CD97 isoforms can bind to integrins  $\alpha 5 \beta 1$  and  $\alpha v \beta 3$ <sup>[22]</sup>; the GAIN domain also binds to CD90 (Thy1), which is expressed by activated endothelial cells<sup>[23]</sup>.

CD97 has also been identified as a tumor-associated receptor as it is significantly up-regulated in many carcinomas, including gastric, colorectal, pancreatic, and other tumors<sup>[24-26]</sup>. Some previous studies indicated that CD97 plays an important role in modulating cell migration and invasion<sup>[27-32]</sup>. Liu *et al*<sup>[31]</sup> reported that two different CD97 isoforms have inverse functions in promoting or suppressing cell migration in gastric cancer. Lin and co-workers<sup>[32]</sup> demonstrated that the overexpression of CD97 on HT1080 fibrosarcoma cells could enhance cell adhesion but reduce cell migration and invasion both *in vitro* and *in vivo*. However, these results were controversial and reflect the complexity of CD97 structure and function. The detailed molecular basis of CD97 in cell adhesion remains to be fully elucidated.

N-glycosylation in CD97 plays an important role in modulating its biological function. A total of 9 potential N-glycosylation sites exist on CD97ECD, including 4 sites in EGF domains and 5 sites in the GAIN domain. Specific N-glycosylation in the GAIN domain regulates GPS auto-proteolysis<sup>[33]</sup>, while N-glycosylation in the EGF domains also influences CD97<sup>EGF</sup> mAb recognition and CD55 ligand binding<sup>[34]</sup>. Furthermore, CD97 is partially N-glycosylated when expressed in smooth muscle cell (SMC) tumors but is not N-glycosylated at all in normal SMC<sup>[14]</sup>; this feature was shown to be a critical factor in the responses of SMC to physiological stimuli.

GPS cleavage of adhesion GPCR yields a heterodimer composed of an extracellular N-terminal fragment (NTF) and a transmembrane C-terminal fragment (CTF)<sup>[35-37]</sup>. The crystal structures of CL1 (CIRL/Latrophilin 1) and BAI3 (brain angiogenesis inhibitor 3) extracellular fragments show why these two fragments can still form a non-covalent heterodimer on the cell surface after proteolysis<sup>[7]</sup>. In these structures, the GPS motif forms a beta-strand and inserts into the hole of the beta-sheet that is formed by the GAIN domain. Strong hydrophobic interactions between these two domains make the GPS

motif unable to disassociate from the GAIN domain unless a tensile force is applied upon ligand binding. Although the structures of EGF domains are also available<sup>[20]</sup>, these independent GAIN or EGF domain structures did not reveal how these tandem domains interact with each other. Specifically, it is not yet known whether these two types of domains can interact with each other and how this may contribute to either their functions or cell adhesion activities. Furthermore, it remains to be determined whether heavy N-glycosylation is involved in the overall conformation of CD97 or CD97-involved cell adhesion. To address these questions, we purified and characterized different variants of CD97ECDs and studied their interactions with HeLa cells. Our data revealed that the EGF5 and GAIN domains are both crucial for HeLa cell binding, while N-glycosylation on the GAIN domain, GPS autoproteolysis and heterodimerization are less important for CD97ECD-HeLa cell interactions. Our data also reveal a potential interface between the GAIN and EGF5 domains as well as the novel concentration-dependent auto-proteolysis of CD97ECD.

## Materials and methods

### Materials

All cell lines were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China) and were originally obtained from the American Type Culture Collection (ATCC). Plasmid pLEXm-His<sub>6</sub> is a custom vector based on pLEXm<sup>[38]</sup>, and pFUSE-hIgG1-Fc2 was purchased from Invitrogen. Ribonuclease A (RA), Ribonuclease B (RB), Ovalbumin (OVA), chondroitin sulfate (CS), and Biotin-LC-NHS were purchased from Sigma. PNGase F was expressed and purified according to a previously published procedure<sup>[39]</sup>.

### Cell culture

A suspension cell line, FreeStyle 293-F (293-F) cells, was cultured with FreeStyle 293 Expression Medium (Gibco, USA) without serum and maintained in a 130 rounds per minute rotating environment. All other culture media were supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 2 mmol/L *L*-glutamine (Gibco, USA). Human embryonic kidney (HEK) 293 cells and a fast-growing variant of HEK293 cells (HEK293FT) were cultured with Dulbecco's minimum essential medium (DMEM; Hyclone, USA). HeLa, hepatocellular carcinoma G2 (HepG2) cells and Chinese hamster ovary (CHO)-K1 cells were cultured in Roswell Park Memorial Institute-1640 (RPMI-1640) medium (Hyclone, USA). All cells were maintained in an incubator with 5% CO<sub>2</sub> and 95% humidity at 37°C.

### Construction of expression vectors

An optimized cDNA sequence of human CD97 was synthesized by GENEWIZ (Suzhou, China), and the sequence of the CD97 extracellular domain is provided as supplementary data. The wild type (WT), EGF-like domains (EGF1-4, EGF1-5) and stalk region (GAIN) of the human CD97 extracellular domain (CD97ECD) were amplified by splicing overlap extension

(SOE) polymerase chain reaction (PCR) using TransStart Fast-Pfu DNA polymerase (TransGen, China), pUC57-CD97 template and the corresponding primers (Supplementary Table S1). The reverse primers of EGF1-4 and EGF1-5 contain an alanine (A) spacer sequence. All Fc-tagged CD97ECD WT and truncated forms were constructed using a secretion expression vector pFUSE-hIgG1-Fc2 (Invitrogen, USA) via *EcoR* I and *Nco* I sites; His<sub>6</sub>-tagged CD97ECD WT was constructed on vector pLEXm-His<sub>6</sub> via a *BsmB* I site. In brief, exacts of PCR products of target genes were incubated along with linear vectors digested by restriction enzyme *EcoR* I, *Nco* I, or *BsmB* I according to the protocol appropriate for a ClonExpress II kit (Vazyme, China). Fc-tagged CD97ECD-S531G mutant and single N-glycosylation site mutants were generated by SOE PCR using vectors that contained CD97ECD WT genes (Fc-tagged CD97ECD WT) as the first template, followed by *Dpn* I (NEB, USA) digestion to eliminate the original template. All constructs were confirmed using DNA sequencing.

#### Transfection and protein purification

CD97ECD expression constructs and a pFUSE-hIgG1-Fc2 empty vector were transfected into HEK293F cells using polyethylenimine (PEI, Sigma, USA). After 72 h of cell culture, conditioned medium containing soluble recombinant proteins was collected by centrifugation. His<sub>6</sub>-tagged proteins were purified using an affinity column that contained Ni-NTA resins (Qiagen, Germany); Fc-tagged proteins were purified using an affinity column that contained Protein A resins (Sigma, USA). All purified proteins were further separated using gel filtration with a Sephadex G-100 (Sigma, USA) column eluted with a 20 mmol/L Tris-HCl (pH 7.4) buffer that contained 5 mmol/L CaCl<sub>2</sub> and 150 mmol/L NaCl. Fractions that contained target proteins were concentrated using 30 kDa centrifugal ultrafiltration tubes (Millipore, USA). Protein concentrations were determined using a BCA protein quantification kit (Beyotime, China).

#### In vitro GPS auto-proteolysis of CD97ECD proteins

Purified Fc-CD97ECD WT protein (0.1 mg/mL) was stored at 4°C or incubated with His-CD97ECD WT protein (2 mg/mL). To analyze the hydrolysis efficiency of Fc-CD97ECD WT, samples were assessed using Western blot with anti-Fc mAb (Sigma) at various time points. The protein band intensity of un-cleaved Fc-CD97ECD WT was quantified by Quantity One 1-D software (Bio-Rad); samples at 0 h was set as a standard.

#### Biotinylation of standard proteins and glycoproteins

To ensure the specific adhesion of CD97ECD with target cells, two glycoproteins, Ribonuclease B (RB) and ovalbumin (OVA), and two standard proteins, Ribonuclease A (RA) and bovine serum albumin (BSA), were labeled with biotin to assess non-specific interactions with these cells. Biotin-LC-NHS was added to the proteins (RA, RB, BSA, and OVA; 1 mg/mL) at a 3:1 molar ratio (Biotin-LC-NHS:Proteins) in 10 mmol/L phosphate buffer (pH 7.4). After incubation at 4°C for 1 h, the solution was subjected to gel filtration separation through a

Sephadex G-25 column (GE, USA) and eluted with 20 mmol/L Tris-HCl (pH 7.4) buffer containing 5 mmol/L CaCl<sub>2</sub> and 150 mmol/L NaCl. Biotinylated protein fractions were combined and concentrated using 10 kDa centrifugal ultrafiltration tubes (Millipore, USA).

#### PNGase F treatment

CD97ECD WT, truncated forms, and mutant versions were treated with PNGase F for deglycosylation. In general, a solution of purified CD97ECD protein and PNGase F in 10 mmol/L phosphate buffer (pH 7.4) was incubated at 4°C for 16 h until the SDS-PAGE analysis revealed no further “downward” shift of the glycoprotein bands. To remove His<sub>6</sub>-tagged PNGase F after deglycosylation, deglycosylated Fc-tagged CD97ECD proteins were purified with an affinity column of Ni-NTA resin (Qiagen) and deglycosylated His<sub>6</sub>-tagged CD97ECD protein was separated using gel filtration with a Sephadex G-100 column, as described above.

#### Western blotting

Purified proteins were subjected to SDS-PAGE separation, then protein bands were transferred to a 0.2 μm PVDF membrane (Millipore, USA) and labeled using rabbit anti-human IgG Fc mAb (Sigma, USA) or mouse anti-His tag mAb (ZSGB-BIO, China), followed by incubation with horseradish peroxidase (HRP)-conjugated goat anti-rabbit or goat anti-mouse IgG secondary Abs (Yeasen, China) for enhanced chemiluminescence (ECL) substrate detection (Pierce, USA) on an ImageQuant LAS 4000 instrument (GE, USA).

#### Mass spectrometry detection

Deglycosylated His<sub>6</sub>-tagged CD97ECD WT treated with PNGase F was resuspended in 100 μL of digestion buffer (100 mmol/L NH<sub>4</sub>HCO<sub>3</sub>, pH 8.0) and digested with trypsin (Sigma, USA) for 16 h. Tryptic peptides were dried using a SpeedVac and dissolved in 4 μL of HPLC buffer A (0.1% formic acid in water, *v/v*) and delivered onto a capillary RPLC trap column through an auto-sampler at a maximum pressure of 250 bars in 100% buffer A. After loading and washing, peptides were transferred to a capillary column connected to an EASY-nLC 1000 HPLC system (Thermo Fisher Scientific, USA). Peptides were eluted with a 70-min gradient of 7% to 80% HPLC buffer B (0.1% formic acid in acetonitrile, *v/v*) in buffer A at a flow rate of 300 nL/min. Eluted peptides were ionized and introduced into a Q Exactive mass spectrometer (Thermo Fisher Scientific, USA) using a nanospray source. A survey of full-scan MS spectra (from *m/z* 350 to 1300) were acquired in an Orbitrap with resolution *r*=70000 at *m/z* 200. In addition, all HPLC/MS/MS data were analyzed using Mascot software (v2.3, Matrix Science Ltd, London, UK). Peak lists were generated with Proteome Discoverer software (version 1.4, Thermo Fisher). Precursor mass tolerance for Mascot analysis was set at ±10 ppm, while fragment mass tolerance was set at ±0.02 Da. The protease was set to trypsin/P, allowing for a maximum of two missed cleavage sites. The data were queried against His<sub>6</sub>-tagged CD97ECD WT protein sequences from

uniprot, including the fixed modification of cysteine residues by carbamidomethylation along with the following variable modifications: oxidation (M), Asn→Asp. All spectra with a Mascot ion score of more than 20 were manually inspected.

#### Cell attachment assay

All cells used for attachment assays were harvested and washed twice with Hanks balanced salt solution (HBSS) that contained 1.7 mmol/L  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  (but not phenol red), pH 7.4. Purified protein (5  $\mu\text{g}$ ) in HBSS buffer (50  $\mu\text{L}$ ) was added to cell suspensions in 96-well plates (Nunc, USA) that were seeded at  $8 \times 10^5$  cells/well. After incubation at 4 °C for 1 h, unbound proteins were washed with HBSS buffer, then blocked with 1% BSA for 30 min. For Fc-tagged CD97ECD samples, cells were probed with rabbit anti-human IgG Fc mAb (Sigma, USA) at 4 °C for 1 h. After extensive washing, cells were mixed with fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit IgG secondary Ab (Yeasen, China) at 4 °C for 40 min. For biotinylated standard proteins, fluorescein isothiocyanate (FITC)-conjugated streptavidin (Yeasen, China) was used for direct detection. After two washes to remove redundant Ab, samples were subjected to flow cytometry analysis (FACSCalibur, BD Bioscience, USA) and data were analyzed using CellQuest software. Further analysis was conducted using WinMDI (version 2.9) and GraphPad Prism (version 6) software packages.

#### Effects of inhibition on cell attachment

Two reagents, chondroitin sulfate (CS) and 5 mmol/L EDTA were used to test for inhibitory activity against CD97ECD-HeLa cell binding. Briefly, purified CD97 protein (5  $\mu\text{g}$ ) was pre-incubated with a HBSS buffer containing CS (5 mg/mL) and/or EDTA (5 mmol/L) in a total volume of 50  $\mu\text{L}$  at 4 °C for 10 min. Then, the mixture was subject to HeLa cell attachment assays, as described above.

#### Small angle X-ray scattering (SAXS) data

After passing through a Sephadex G-200 column (GE, USA) to avoid protein aggregation in the sample solutions, His<sub>6</sub>-CD97ECD WT and its deglycosylated form were subjected to SAXS analysis on a BRUKER SAXS instrument at the BL19U2 Bio-SAXS beamline of the Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, China). The photon energy of X-rays range from 7 keV to 15 keV and the luminous flux in samples was at least  $4 \times 10^{12}$  phs/s. The focusing spot size was less than  $380 \times 110 \mu\text{m}^2$ . All data collection was carried out at 4 °C. Additionally, 20 mmol/L Tris-HCl (pH 7.4) buffer containing 5 mmol/L  $\text{CaCl}_2$  and 150 mmol/L NaCl was used for protein solutions (1.5 mg/mL) and as a blank control. The 2D scattering images were converted by Fit2D software into 1D data, and then PRIMUS software was used for calculations. All sample scatterings had the solvent background scattering deducted from the measured values.

#### Molecular modeling

In parallel with the *ab initio* methods used to construct low-

resolution envelopes from the SAXS data, homology modeling and rigid-body docking were performed to obtain atomic resolution models for both His<sub>6</sub>-tagged CD97ECD WT and its deglycosylated form. First, five homology models of the EGF 1–5 domains of CD97ECD were generated using the SWISS-MODEL server<sup>[40]</sup> based on the crystal structure of EGF domains 1, 2, and 5 of human adhesion GPCR E2 (EMR2) (PDB ID: 2BO2, identity scores of 95%, 98%, 53%, 65%, and 100% for EGF 1–5 domains, respectively) as a template<sup>[20]</sup>. The GAIN domain homology model was obtained using the LOM-ETS serve<sup>[41]</sup> based on the crystal structures of the GAIN and HormR domains of brain angiogenesis inhibitor 3 (BAI3) (PDB ID: 4DLO, identity score of 18%) as templates<sup>[7]</sup>. These models were then used for rigid-body docking. Rigid-body docking of the structures of individual domains of CD97ECD into two dummy atom models were performed manually based on the shape and connections between the C- and N-termini of adjacent fragments.

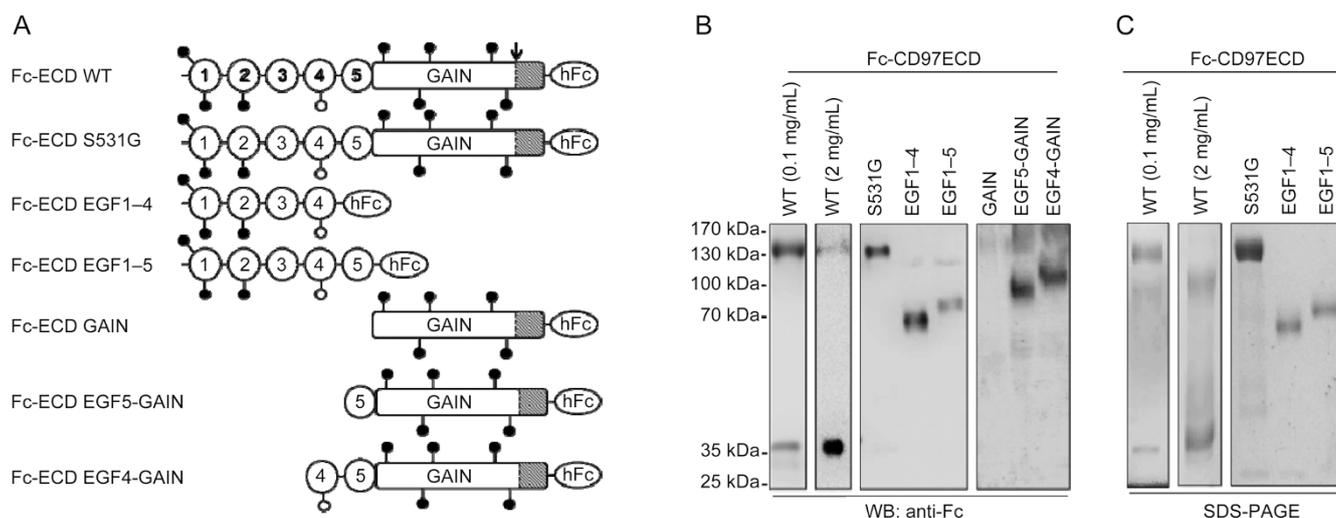
## Results

### Expression of CD97 extracellular fragments

We tested the secretion of CD97 fragments from HEK293F transfected cells with C-terminal Fc tags (Figure 1A). While the construct ending at L530 (GPS cleavage occurred between L530 and S531) yielded no expression (data not shown), the full extracellular domains (both wild-type and the auto-proteolysis mutant S531G) were expressed robustly (5–6 mg/L) when the C-termini were extended to E543 (Figure 1B). This finding is in accord with previous structures showing that the GPS motif is an integral part of the GAIN domain. While EGF1-4 (21–209) was expressed at a similar level as full ECDs, the expression level of EGF1-5 (21–258) decreased significantly (Figure 1B). Furthermore, the “GAIN domain only” construct (residues 262–543) was not secreted at all (Figure 1B). Thus, we hypothesized that potential interactions occur between the EGF5 and GAIN domains and the EGF5-GAIN complex may be important for proper folding and expression. We tested this hypothesis by expressing two additional constructs starting at EGF4 or EGF5 and extending to the C-terminus of the GAIN domain. Our findings confirmed our hypothesis and these two constructs were robustly expressed (Figure 1B). Our SAXS data are also agreement with our expression results (see Results and Discussion sections below).

### CD97 GPS auto-proteolysis is concentration-dependent and regulated by N-glycosylation

We next studied the GPS cleavage of these proteins using SDS-PAGE and Western blotting (Figure 1B). Interestingly, we observed concentration-dependent hydrolysis at the GPS site. WT CD97ECD at 0.1 mg/mL displayed a major band at approximately 130 kDa, which matched the size of non-cleaved CD97ECD plus the Fc tag (CD97ECD and the Fc are 100 and 35 kDa, respectively). Anti-Fc Western blotting was then used to differentiate the Fc-containing fragments from other fragments. At a concentration of 0.1 mg/mL, undigested CD97ECD-Fc was the main form (approximately 70%–80%



**Figure 1.** Truncated and GPS site mutated forms of Fc-CD97 ECD. (A) Schematic overview of CD97ECD-hFc fusion proteins. Circle, EGF-like domain; Rectangle, GAIN domain; Lollipop, potential N-glycosylation sites; Arrow, GPS auto-proteolytic site. (B and C) Purified Fc-CD97ECD fragments analyzed by Western blot (B) or Coomassie brilliant blue staining (C). Blots were probed with anti-human IgG Fc mAb, followed by horseradish peroxidase (HRP)-conjugated Ab.

estimated from the gel) and the percentage remained similar after incubation at 4°C for 48 h (Figure 2A and Supplementary Figure S3), suggesting that the auto-proteolysis process occurs very slowly. However, when CD97ECD-Fc was concentrated to 2 mg/mL, the rate of auto-proteolysis dramatically increased. Indeed, the undigested form was not visible using SDS-PAGE, and only a weak band was identified in the anti-Fc lane (Figure 1B). Concentration-dependent auto-proteolysis suggests an intermolecular mechanism operating for the CD97 EGF1-5 isoform rather than the intramolecular mechanism that was reported for other CD97 isoforms<sup>[33, 35]</sup>. We then added a high-concentration (2 mg/mL) sample of His-tagged CD97ECD into a low-concentration (0.1 mg/mL) Fc-tagged CD97ECD sample at 4°C, and then monitored the auto-proteolysis rate of CD97ECD-Fc using anti-Fc Western blotting; only the low-concentration Fc-tagged CD97 protein band was observed in the blot (Figure 2A). Compared with the controls, CD97ECD-Fc auto-proteolysis was significantly accelerated. We further tested this phenomenon by titrating four different concentrations of His-tagged CD97ECD, and our results suggested that cleavage is apparently greater at a higher concentration (4 mg/mL) than for low concentration samples, especially over the 48 h-period (Supplementary Figure S3). This finding supports a hypothesis of intermolecular proteolysis. No auto-proteolysis was detected in the GPS site mutant S531G or GPS truncated forms EGF1-4 and EGF1-5 (Figure 1B).

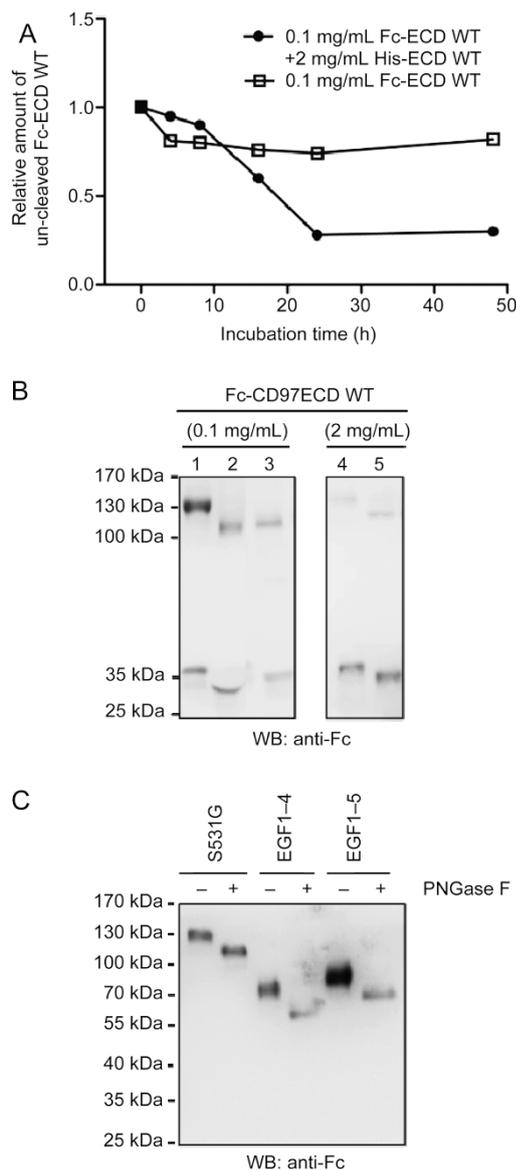
PNGase F catalyzed deglycosylation of WT CD97ECD under 0.1 mg/mL and reduced the protein band size to approximately 110 kDa; furthermore, its GPS cleaved-to-uncleaved ratio remained largely unchanged (Figure 2B, Lane 2). Thereafter, the deglycosylated form was concentrated to 2 mg/mL. In contrast to WT CD97ECD, auto-proteolysis ceased at a

level similar to that of the low concentration sample (Figure 2B, Lane 3). This finding demonstrates that in addition to the protein concentration, N-glycosylation also regulates the GPS cleavage rate of CD97 EGF1-5, similarly to previous reports for CD97 EGF1,2,5<sup>[33]</sup>.

We then studied the effects of specific N-glycosylation sites on GPS auto-proteolysis (Figure 3). Three representative single mutants were selected based on the locations of the sites. N203 is located in the fourth EGF-like repeat, a position that is involved in CS binding<sup>[10]</sup>. N371 is only one residue away from the RGD integrin-binding motif<sup>[22]</sup>. N453 is located close to the GPS motif; however, it is not conserved and is present in *Homo sapiens* and *Chlorocebus* but not in *Canis*, *Ursus*, *Bos*, *Rattus* or *Cricetulus*. Because the auto-proteolysis of CD97 is predominantly regulated by the GAIN domain<sup>[33]</sup>, we also designed double- and quintuple-mutants within this domain (N406/N413, N453/N520 and N371/N406/N413/N453/N520) that are analogous to those generated in a previous study<sup>[33]</sup>. We found that the single mutants N203, N371, and N453 behaved like WT CD97ECD, with less GPS cleavage at a lower concentration (0.1 mg/mL) and robust cleavage at a higher concentration (2 mg/mL) (Figure 3B). Mutations at multiple N-glycosylation sites differentially affected the GPS hydrolytic reaction (Figure 3C). Double mutation at N406/N413 suppressed auto-proteolysis, while the N371/N453 mutant did not influence auto-proteolysis and N453/N520 slightly reduced auto-proteolysis similar to the GPS cleaved fragment, while at a higher concentration it increased less than the WT. The quintuple mutant (N371/N406/N413/N453/N520) completely lost GPS proteolytic activity.

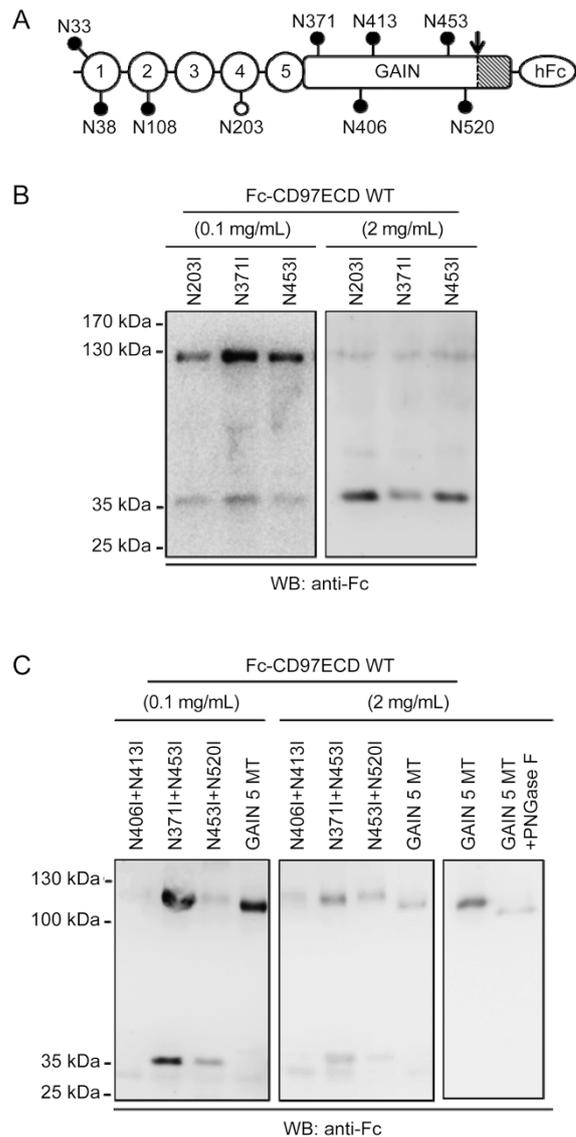
#### Mass spectrometry identification of deglycosylated CD97ECD

In addition to Fc-tagged fragments, we also expressed His<sub>6</sub>-



**Figure 2.** *In vitro* GPS auto-proteolysis is concentration-dependent and regulated by N-glycosylation. (A) Time-course profiles on auto-proteolysis of Fc-CD97ECD WT at 0.1 mg/mL and the mixture of Fc-CD97ECD WT (0.1 mg/mL) and a high-concentration (2 mg/mL) sample of His-CD97ECD WT. The samples at 0 h was set as standard. (B) Western blot analysis of Fc-CD97ECD WT after PNGase F treatment. Lane 1: Fc-CD97ECD WT (0.1 mg/mL); lane 2: Fc-CD97ECD WT (0.1 mg/mL) after PNGase F treatment; lane 3: Fc-CD97ECD WT (0.1 mg/mL) after PNGase F treatment then concentrated to 2 mg/mL; lane 4: Fc-CD97ECD WT (2 mg/mL); lane 5: Fc-CD97ECD WT (2 mg/mL) after PNGase F treatment. (C) Western blot analysis of Fc-CD97ECD S531G, EGF1-4, EGF1-5 and their deglycosylated forms by PNGase F treatment. All blots were probed by anti-Fc mAb.

tagged CD97ECD. Purified CD97ECD-His was subjected to deglycosylation by PNGase F and then was verified using mass spectrometry. PNGase F treatment removed the carbohydrates and also changed Asn residues to Asp, resulting in a 1 Dalton difference in molecular weight. After trypsin



**Figure 3.** N-Glycosylation mutants of Fc-CD97ECD and GPS auto-proteolysis. (A) Diagram representation shows the N-glycosylation site information of Fc-CD97ECD WT. Solid lollipop, N-glycosylation sites (identified by MS); Hollow lollipop, potential N-glycosylation site which was not detected by MS. (B and C) Western blot analysis of single, double, or multiple N-glycosylation site mutants of Fc-CD97ECD. GAIN 5 MT is the quintuple mutant (N371I/N406I/N413I/N453I/N520I).

digestion, our peptide mapping results showed that eight out of nine N-glycosylation sites had been deglycosylated (Table 1, Supplementary Figure S5). Conversion of N203 to Asp was not detected, possibly because of the absence of glycosylation or steric inaccessibility to the enzyme. GPS auto-proteolysis was also confirmed as the peptide containing L530 (Table 1, bottom row) was observed as being hydrolyzed from S531.

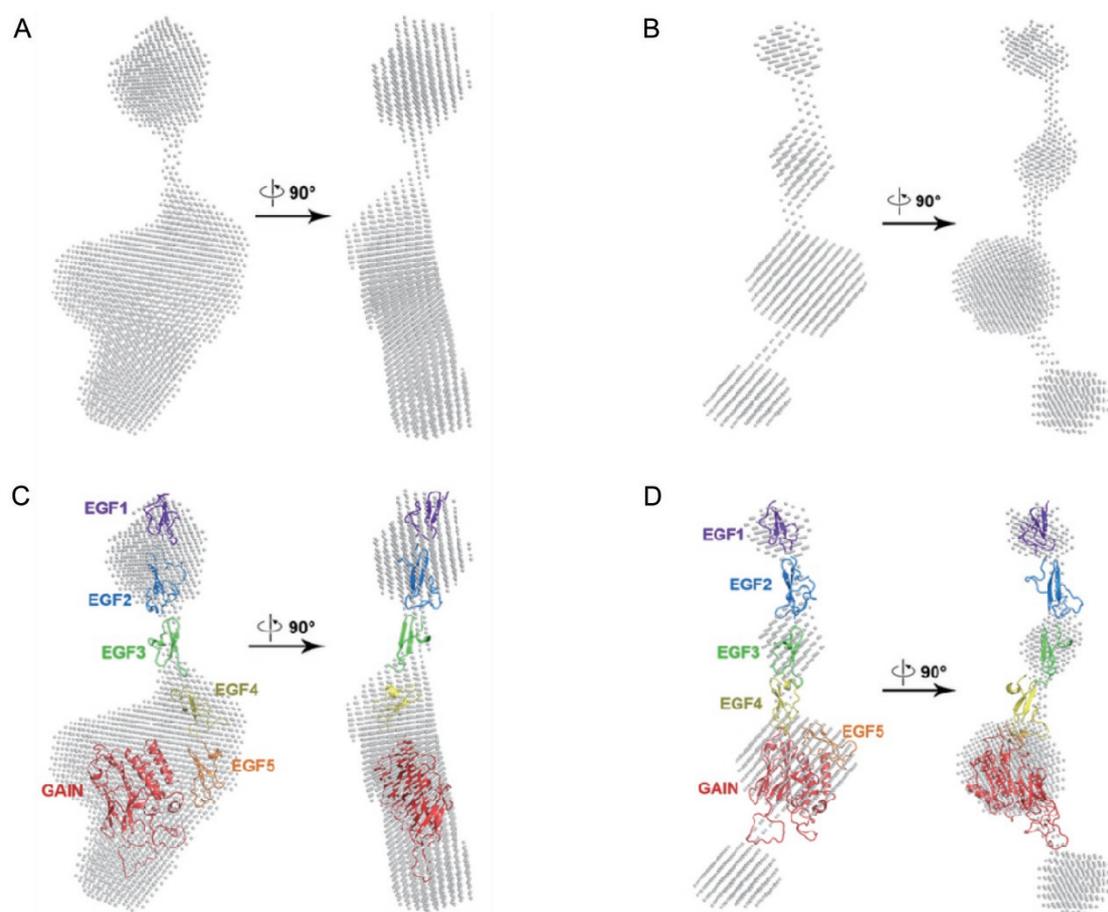
**SAXS measurements of the CD97ECD WT and PNGase F-digested forms**  
WT CD97ECD and its deglycosylated form were then sub-

jected to SAXS measurements. As shown in Figure 4A, the overall shape of WT CD97ECD appears like a spatula with the GAIN domain representing the body and the EGF domains constituting the handle. The profile of the deglycosylated form is relatively thin, which is consistent with its lack of carbohydrates (Figure 4B). We have generated models of the CD97ECD GAIN and EGF domains and docked them into SAXS envelopes in a rigid docking mode. Homology models of the GAIN and EGF1-5 domains were derived based on structures of BAI3 (GAIN and HormR domains) and EMR2 (EGF domains 1, 2, 5)<sup>[7, 20]</sup>. The handles of the spatulas were not long enough, and we could only fit 3 or 4 of the EGF domains into the WT and deglycosylated forms, respectively, suggesting that at least the EGF5 domain might co-localize with the GAIN domain in the large blob or body of the spatula (Figure 4C and 4D). The blobs in both forms are large enough to accommodate both the GAIN and EGF5 domains. In agreement with our previous expression findings, this phenomenon also suggests that an interface exists between these two domains, which could be the reason that both the EGF5-containing construct and GAIN domain cannot be expressed by themselves. Although PNGase F treatment removed car-

**Table 1.** N-Glycosylation site identification of His-CD97ECD by MS determination.

MS detected peptide	N-glycosylation site
WCPQNSSCVNATACR	N33/38→D33/38
WCPQNSSCVNATACR/ GCARWCPQNSSCVNATACR	N38→D38
TFKNESENTCQDVDECQQNPR	N108→D108
GDKNVTMGQSSAR	N371→D371
LNWAVAAGAEDPGPAVAGILSI QNMPTLLANASLNLHLSK	N406/413→D406/413
RLSAVNSIFLSHNNTK	N453→D453
NGSTTCQCSSL	N520→D520

bohydrates, the standard deviations of Dmax of the deglycosylated form (204 Å) was approximately 7 Å longer than that of WT CD97ECD, suggesting that a potential orientational change occurs at some point after digestion (Figure 4). Obviously, elucidating the molecular details, including the unexplained small blob at the bottom of each protein form, will



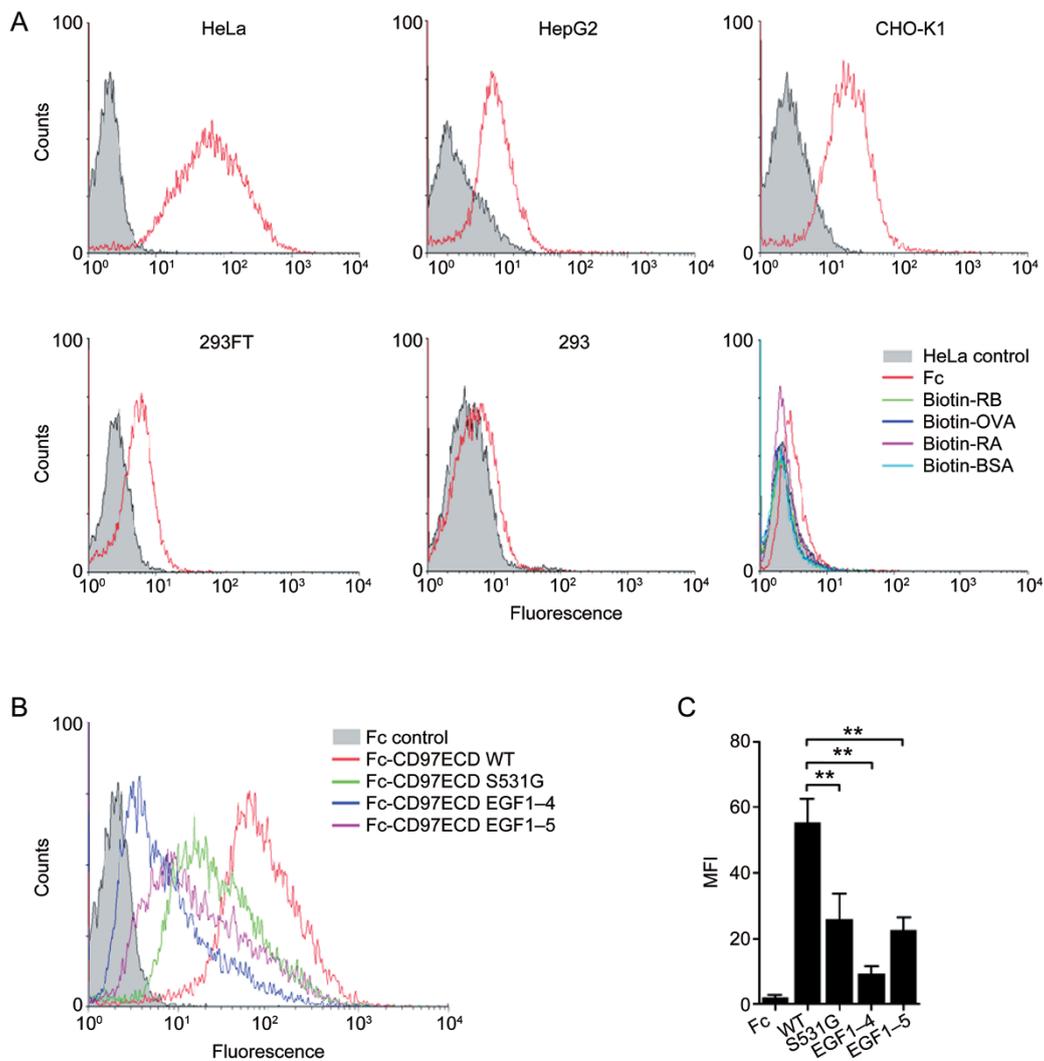
**Figure 4.** PNGase F deglycosylation of CD97ECD caused shape change in SAXS measurement. (A) SAXS envelope of His-CD97ECD WT protein shape. (B) SAXS envelope of PNGase F digested His-CD97ECD WT protein shape. (C) Docking results of CD97ECD homology models to SAXS envelope of His-CD97ECD WT. (D) Docking results of CD97ECD homology models to SAXS protein shape of PNGase F digested His-CD97ECD WT.

require a high-resolution crystal structure to be solved.

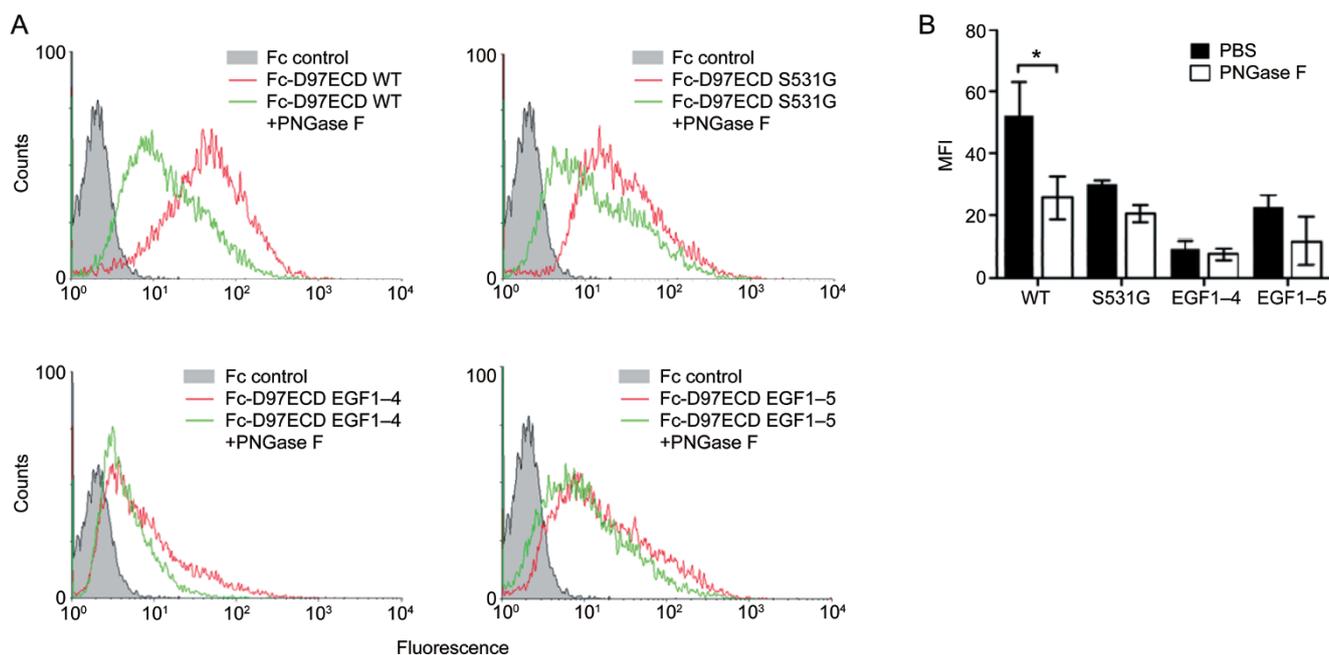
#### Molecular basis of CD97ECD mediated HeLa cell attachment

We sought to investigate the molecular details of CD97-mediated cell attachment using the various CD97ECD forms prepared above (Figure 1C). We screened several cell lines in CD97 attachment assays, as shown in Figure 5. Flow cytometry data indicated that WT CD97ECD bound to the cervical cancer HeLa cell line more strongly than other cell lines that we tested. Indeed, HepG2 and CHO-K1 cells could also interact with CD97, while HEK293 and HEK293FT bound poorly to WT CD97ECD. Therefore, we chose HeLa cells as the model cells to examine other variants in CD97. As controls, we also measured five unrelated proteins (RA, RB, OVA, BSA, and Fc) and found that none of them showed a significant interaction with HeLa cells (Figure 5A). Inhibition experiments using CS and/or EDTA dramatically reduced CD97-HeLa cell attach-

ment, suggesting that the binding probably was a consequence of recognition of CD97 and CS on the cell surface that occurred in a calcium-dependent manner<sup>[10]</sup> (Supplementary Figure S4). Furthermore, four CD97ECD-Fc fragments showed distinct binding affinities to HeLa cells (Figure 5B and 5C). Compared with the full-length ECD, C-terminal truncations to EGF5 and EGF4 yielded progressively worse binding, indicating that both EGF domains and the GAIN domain are involved in HeLa cell attachment. The mean fluorescence intensity (MFI) of CD97ECD-S531G dropped to approximately half of the WT, and PNGase F treatment of these four CD97 proteins consistently caused reductions in cell attachment (Figure 6). We considered that these data suggest that GPS auto-proteolysis and N-glycosylation of CD97 may be involved in HeLa cell adhesion. However, when we mutated the N-glycosylation sites of CD97 for deglycosylation and auto-proteolysis inhibition, the N-glycosylation mutants showed a nearly similar binding



**Figure 5.** HeLa cell attachment assay for Fc-CD97ECD WT and its GPS site mutant and truncated forms. (A) FACS profile of attachment of Fc-CD97ECD WT with several cell lines. Biotin-RA/RB/OVA/BSA and Fc fragment marginally attached to HeLa cells (bottom right panel). (B) FACS profile of HeLa cell attachment with four Fc-CD97ECD proteins. (C) Mean fluorescence intense (MFI) data of HeLa cell attachment with four Fc-CD97ECD proteins. Quantification from three independent experiments: mean±SD. \*\**P*<0.01, two tailed unpaired Student's *t* test.



**Figure 6.** HeLa cell attachment assay for PNGase F digested Fc-CD97ECD proteins. (A) FACS profiles of HeLa cell attachment with Fc-CD97ECD proteins and their deglycosylated forms by PNGase F digestion. (B) Mean fluorescence intense (MFI) data of HeLa cell attachment with Fc-CD97ECD proteins and their deglycosylated forms by PNGase F digestion. Quantification from three independent experiments: mean $\pm$ SD. \* $P$ <0.05, two tailed unpaired Student's  $t$  test.

affinity as the WT, which argues for a different mechanism (see Results and Discussion below).

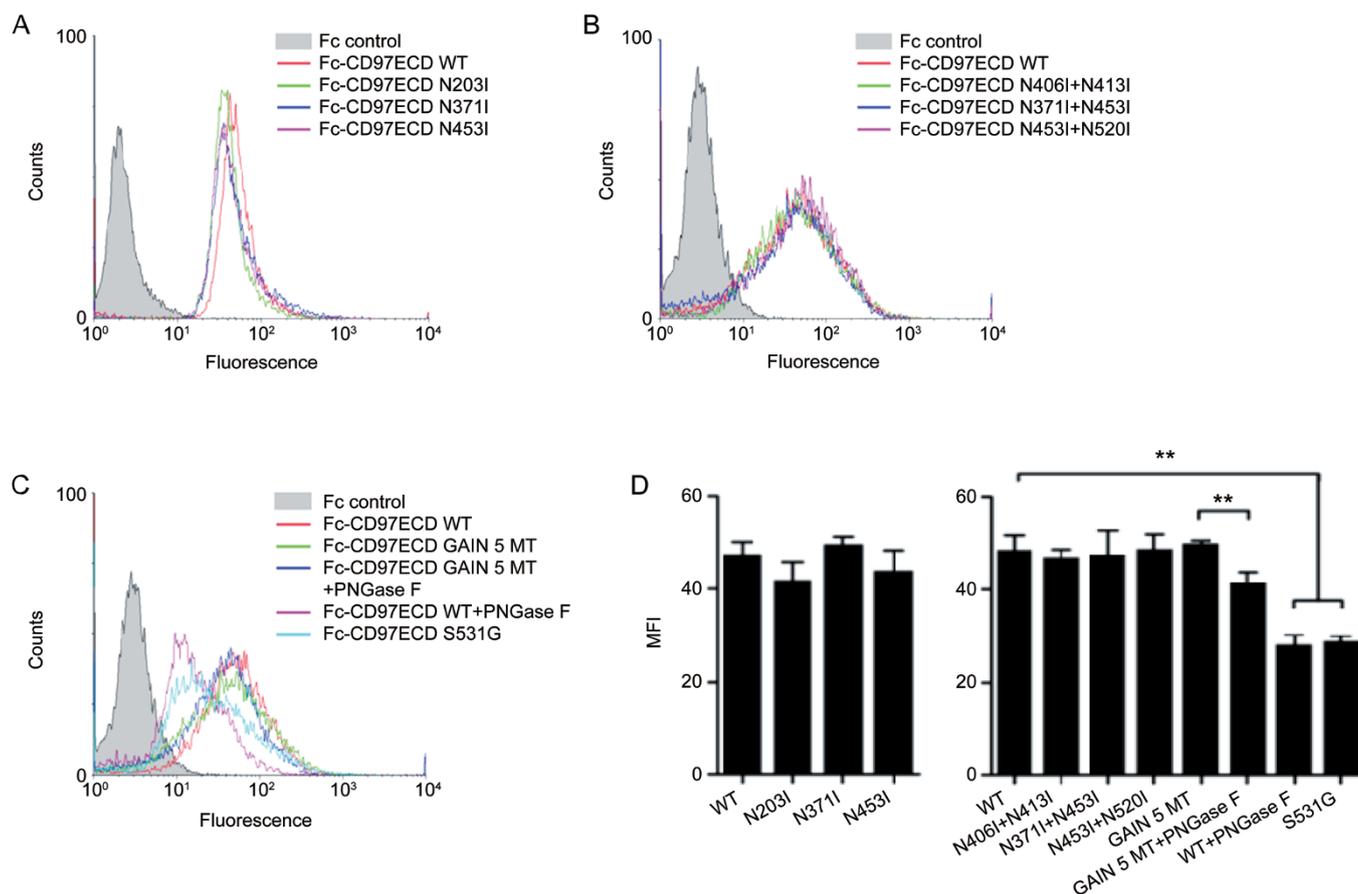
#### N-glycosylation of the CD97 GAIN domain and GPS auto-proteolysis are not required for HeLa cell attachment

PNGase F non-specifically removes all carbohydrates, and to better refine our data, we wanted to study the function of specific N-glycosylation sites on the GAIN domain of CD97 by making specific mutants (Figure 7). Surprisingly, all of the mutants that we generated—the single mutants (N203I, N371I, N453I), double mutants (N406I/N413I, N371I/N453I, N453I/N520I), and quintuple mutant (N371I/N406I/N413I/N453I/N520I)—maintained almost identical HeLa binding activity as WT CD97ECD. Further deglycosylation by PNGase F on the quintuple mutant slightly reduced cell binding, but the reduction was small and less significant compared with PNGase F-treated CD97 WT. These data illustrate that N-glycosylation of the GAIN domain plays a minor role in CD97-mediated attachment to HeLa cells. Together, the results shown in Figures 6 and 7 indicate that the N-glycosylation of EGF domains may more strongly contribute to HeLa cell adhesion. Alternatively, possible differences in conformation observed in SAXS and modeling data for PNGase F-deglycosylated CD97 WT, as described above, might explain the greater reduction of its HeLa binding affinity compared with the PNGase F-treated CD97 quintuple mutant. Furthermore, the quintuple mutant completely inhibited GPS auto-proteolysis (Figure 3C), while its HeLa cell binding activity remained unchanged (Figure 7C). These data indicate that GPS cleavage is not required

for HeLa cell attachment, which is contradictory to the HeLa attachment findings made using the S531G mutant. A possible explanation is that the GPS-mutated variant of CD97 may somehow have altered local structures or biological features. Arriving at a more detailed structural mechanism will require CD97 crystal structures to be solved.

#### Discussion

In this present study, we report the successful expression of CD97ECD and EGF domains in HEK293F cells. Compared with the glycosylation patterns obtained using bac-to-bac or other expression systems, that of proteins expressed by a mammalian cell line represents a more physiological approach to study CD97 adhesion<sup>[42]</sup>. Although the EGF1-4 fragment can be expressed at the same level as WT CD97ECD, the reduced expression of EGF1-5 and lack of secretion of the GAIN domain indicate that the EGF5 and GAIN domains may directly interact with each other. Perhaps notably, in the previously reported structures of CL1 and BAI3, the GAIN domains were co-crystallized with the hormone-binding domain (HormR) and the interface between the HormR domain and GAIN subdomain A is very well conserved<sup>[7]</sup>. For CD97, the chaperone of the GAIN domain might be the EGF5 domain. Our SAXS data for both the WT and deglycosylated samples suggest that a potential interface exists between the GAIN and EGF5 domains. Additionally, the importance of EGF5 may also be indicated by its presence in all CD97 isoforms<sup>[18]</sup>. The other two EGF domains that are present in all isoforms are EGF 1 and 2; the importance of these isoforms



**Figure 7.** GPS auto-proteolysis and N-glycosylation are not required for HeLa cell attachment. (A–C) FACS profiles of HeLa cell attachment with N-glycosylation site mutants of Fc-CD97ECD. (D) Mean fluorescence intensity (MFI) data of HeLa cell attachment with N-glycosylation site mutants of Fc-CD97ECD. Quantification from three independent experiments: mean±SD. \*\**P*<0.01, two tailed unpaired Student's *t* test.

is illustrated by their involvement in the interaction between CD97 and its counterpart CD55<sup>[20]</sup>. By contrast, to the best of our knowledge, the functions of EGF3 and 4 have not yet been reported in detail.

The role of GPS auto-proteolysis in the activity of adhesion GPCRs has been associated with its biological functions<sup>[43, 44]</sup>. The group of Lin originally investigated GPS auto-proteolysis and reported a self-catalytic reaction of EMR2 via an ester intermediate and also characterized the GPS cleavage site<sup>[35, 36]</sup>. They also elucidated the regulation of GPS auto-proteolysis of the CD97 EGF1,2,5 isomer by specific N-glycosylation patterns<sup>[33]</sup>. EMR2 is highly homologous to CD97, especially on the ECD; however, the GPS auto-cleavage patterns are distinct. While EMR2 auto-proteolysis occurs *in vitro* and cleavage can be accelerated in the presence of NH<sub>2</sub>OH, the auto-proteolysis of CD97(EGF1,2,5) plateaus at approximately 50% *in vitro* and is not affected by incubation with NH<sub>2</sub>OH at 37°C, even after 16 h. In our present report, we focused on the CD97(EGF1-5) isoform and found that GPS auto-proteolysis occurs in a unique manner that differs from that of CD97(EGF1,2,5) and EMR2. Unlike CD97(EGF1,2,5), GPS cleavage of CD97(EGF1-5) can occur *in vitro* in a concentration-dependent

manner. Our data unambiguously indicate that a high concentration dramatically accelerates GPS cleavage. The concentration-dependent auto-proteolysis of CD97(EGF1-5) suggests a mechanism for intermolecular auto-proteolysis, which has also been shown in cases of initiator caspases<sup>[45]</sup>. A possible reason for the difference in proteolysis mechanisms between CD97(EGF1-5) and CD97(EGF1-5) or its homolog protein EMR2 may be the domain organization. While CD97(EGF1-5) has all 5 EGF domains, the other two isoforms lack EGF3 and EGF4. From our SAXS data and modeling studies, EGF4 is located at the neck of the “spatula” between the “handle” of EGF1-3 and “body” of the GAIN domain and EGF5. It is possible that the absence of EGF4 may alter the connection and orientation between the handle and body of CD97, which may further affect auto-proteolysis.

Differences in auto-proteolysis, combined with distinct domain organization and glycosylation status, likely leads to functional diversity. For example, N-glycosylation similarly regulates auto-proteolysis for both the CD97 EGF1-5 and EGF1,2,5 isoforms, except that the N453/N520 double-mutants promote the GPS-mediated hydrolysis of CD97 EGF1,2,5 more aggressively than the EGF1-5 isoform. Previously<sup>[10]</sup>, EMR2

exhibited high binding affinities to several cell lines, including HEK293T, CHO-K1, NIH3T3, COS-7, and HeLa cells. In our present study, CD97ECD selectively binds to HeLa cells much more strongly than the other cell lines that we tested. Aberrant over-expression of CD97 is often observed in various cancers and CD97 can promote cell-cell aggregation by the up-regulation of N-cadherin expression<sup>[46]</sup>. The high affinity binding of CD97 to cancer cells suggests its potential function in tumor cell proliferation and metastasis.

GPS auto-proteolysis is essential for the biological functions of many adhesion GPCRs<sup>[43,44]</sup>. In this present work, we demonstrated that GPS cleavage is not required for CD97-HeLa cell attachments. Inhibition of GPS hydrolysis by CD97ECD was achieved by three distinct approaches—GPS site mutation, PNGase F deglycosylation, and N-glycosylation site mutations. Each of these approaches required manipulation of the WT form that could represent a risk of inducing protein dysfunction. In this present case, PNGase F digestion and mutations of the GPS site reduced cell binding, while N-glycosylation site mutations retained the attachment activity. Logically, the suppression of auto-proteolysis by glycosylation mutants should be irrelevant to the cell attachment. This finding will be helpful and valuable for further functional studies of CD97, especially for studies of GPS auto-proteolysis because conclusions could otherwise be biased if they are only based on a single experimental approach. In addition, it is notable that these data are based on soluble CD97ECD, which may not represent a situation in which CD97 is expressed on the cell surface because a tensile force may be applied to CD97 upon binding to ligand that may trigger release of the ECD from the transmembrane domain via the GPS motif.

PNGase F is a glycosidase that is widely used to remove N-glycans from glycoproteins for studies of the role of glycosylation in protein function<sup>[47]</sup>. It is also extensively used in protein crystallization studies because glycans are usually flexible and can impede crystal formation. Some solved GPCR crystal structures are deglycosylated by PNGase F<sup>[48–50]</sup>. In the present study, the PNGase F digested and N-glycosylation mutated forms of CD97 exhibited functional variation in cell binding activities. Usually, PNGase F cleaves N-glycans of folded glycoproteins with less risk of conformational change, whereas N-glycosylation sites mutagenesis is involved in glycoprotein folding and may result in an increased risk of changes in the conformation or flexibility of glycoproteins. However, our experimental data suggest that deglycosylation could be more complex and exceed our initial predictions. Our use of different approaches to control glycosylation provides a new perspective for functional studies of glycoproteins such as CD97.

Overall, our biochemical and cellular data reveal new features of CD97ECD, including a potential interface between the GAIN and EGF5 domains and a novel concentration-dependent auto-proteolysis mechanism for CD97ECD. Although we have observed possible structural differences between WT CD97ECD and its PNGase F deglycosylated form in SAXS data and modeling results, it is difficult to conclude

that PNGase F treatment induces conformational changes because of the low resolution SAXS data. Further detailed structural studies based on the CD97 crystal structures may help to address this question.

### Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 21372238 and 21572244 to Wei HUANG). Gao-jie SONG was supported by a State Education Ministry grant for Returned Overseas Chinese Scholars (No. X-0801-15-020). We thank the staff at Beamline BL19U2 of the National Center for Protein Science Shanghai (NCPSS) at the Shanghai Synchrotron Radiation Facility (SSRF) for help with SAXS data collection. We would also like to thank colleagues at the iHuman Institute for providing technical support and helpful discussions, and Jack SKINNER for his critical reading of this manuscript.

### Author contribution

Wei HUANG and Gao-jie SONG conceived the ideas for this study and guided the research; Li-yun YANG performed the major experiments; Xiao-fang LIU, Yang YANG, Yu-bo TANG, and Kai-wen LIU participated in the construction of CD97 variants; Lin-lin YANG and Huai-yu YANG performed molecular modeling; Min ZHANG and Min-jia TAN performed mass-spec analysis for N-glycosylation site identification; Shan-mei CHENG and Ye-chun Xu analyzed SAXS data; Zhi-jie LIU, Wei HUANG, and Gao-jie SONG wrote the manuscript.

### Supplementary information

Supplementary information is available on the website of *Acta Pharmacologica Sinica*.

### References

- 1 Attwood TK, Findlay JB. Fingerprinting G-protein-coupled receptors. *Protein Eng* 1994; 7: 195–203.
- 2 Kolakowski LJ. GCRDb: a G-protein-coupled receptor database. *Receptors Channels* 1994; 2: 1–7.
- 3 Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* 2003; 63: 1256–72.
- 4 Stacey M, Lin HH, Gordon S, McKnight AJ. LNB-TM7, a group of seven-transmembrane proteins related to family-B G-protein-coupled receptors. *Trends Biochem Sci* 2000; 25: 284–9.
- 5 Yona S, Lin HH, Siu WO, Gordon S, Stacey M. Adhesion-GPCRs: emerging roles for novel receptors. *Trends Biochem Sci* 2008; 33: 491–500.
- 6 Langenhan T, Aust G, Hamann J. Sticky signaling — adhesion class G protein-coupled receptors take the stage. *Sci Signal* 2013; 6: re3.
- 7 Arac D, Bolliger MF, Nguyen J, Soltis SM, Sudhof TC, Brunger AT. A novel evolutionarily conserved domain of cell-adhesion GPCRs mediates autoproteolysis. *EMBO J* 2012; 31: 1364–78.
- 8 Promel S, Langenhan T, Arac D. Matching structure with function: the GAIN domain of adhesion-GPCR and PKD1-like proteins. *Trends Pharmacol Sci* 2013; 34: 470–8.
- 9 Aust G. Adhesion-GPCRs in tumorigenesis. *Adv Exp Med Biol* 2010;

- 706: 109–20.
- 10 Stacey M, Chang GW, Davies JQ, Kwakkenbos MJ, Sanderson RD, Hamann J, *et al*. The epidermal growth factor-like domains of the human EMR2 receptor mediate cell attachment through chondroitin sulfate glycosaminoglycans. *Blood* 2003; 102: 2916–24.
  - 11 Hamann J, Vogel B, Van-Schijndel GM, Van-Lier RA. The seven-span transmembrane receptor CD97 has a cellular ligand (CD55, DAF). *J Exp Med* 1996; 184: 1185–9.
  - 12 Xu L, Hynes RO. GPR56 and TG2: possible roles in suppression of tumor growth by the microenvironment. *Cell Cycle* 2007; 6: 160–5.
  - 13 Park D, Tosello-Tramont AC, Elliott MR, Lu M, Haney LB, Ma Z, *et al*. BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. *Nature* 2007; 450: 430–4.
  - 14 Aust G, Wandel E, Boltze C, Sittig D, Schutz A, Horn LC, *et al*. Diversity of CD97 in smooth muscle cells. *Cell Tissue Res* 2006; 324: 139–47.
  - 15 Hamann J, Eichler W, Hamann D, Kerstens HM, Poddighe PJ, Hoovers JM, *et al*. Expression cloning and chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretion receptor superfamily with an unusual extracellular domain. *J Immunol* 1995; 155: 1942–50.
  - 16 Jaspars LH, Vos W, Aust G, Van-Lier RA, Hamann J. Tissue distribution of the human CD97 EGF-TM7 receptor. *Tissue Antigens* 2001; 57: 325–31.
  - 17 Eichler W, Hamann J, Aust G. Expression characteristics of the human CD97 antigen. *Tissue Antigens* 1997; 50: 429–38.
  - 18 Gray JX, Haino M, Roth MJ, Maguire JE, Jensen PN, Yarme A, *et al*. CD97 is a processed, seven-transmembrane, heterodimeric receptor associated with inflammation. *J Immunol* 1996; 157: 5438–47.
  - 19 Hamann J, Stortelers C, Kiss-Toth E, Vogel B, Eichler W, Van-Lier RA. Characterization of the CD55 (DAF)-binding site on the seven-span transmembrane receptor CD97. *Eur J Immunol* 1998; 28: 1701–7.
  - 20 Abbott RJ, Spendlove I, Roversi P, Fitzgibbon H, Knott V, Teriete P, *et al*. Structural and functional characterization of a novel T cell receptor co-regulatory protein complex, CD97-CD55. *J Biol Chem* 2007; 282: 22023–32.
  - 21 Kwakkenbos MJ, Pouwels W, Matmati M, Stacey M, Lin HH, Gordon S, *et al*. Expression of the largest CD97 and EMR2 isoforms on leukocytes facilitates a specific interaction with chondroitin sulfate on B cells. *J Leukoc Biol* 2005; 77: 112–9.
  - 22 Wang T, Ward Y, Tian L, Lake R, Guedez L, Stetler-Stevenson WG, *et al*. CD97, an adhesion receptor on inflammatory cells, stimulates angiogenesis through binding integrin counterreceptors on endothelial cells. *Blood* 2005; 105: 2836–44.
  - 23 Wandel E, Saalbach A, Sittig D, Gebhardt C, Aust G. Thy-1 (CD90) is an interacting partner for CD97 on activated endothelial cells. *J Immunol* 2012; 188: 1442–50.
  - 24 Aust G, Eichler W, Laue S, Lehmann I, Heldin NE, Lotz O, *et al*. CD97: a dedifferentiation marker in human thyroid carcinomas. *Cancer Res* 1997; 57: 1798–806.
  - 25 Steinert M, Wobus M, Boltze C, Schütz A, Wahlbuhl M, Hamann J, *et al*. Expression and regulation of CD97 in colorectal carcinoma cell lines and tumor tissues. *Am J Pathol* 2002; 161: 1657–67.
  - 26 Aust G, Steinert M, Schütz A, Boltze C, Wahlbuhl M, Hamann J, *et al*. CD97, but not its closely related EGF-TM7 family member EMR2, is expressed on gastric, pancreatic, and esophageal carcinomas. *Am J Clin Pathol* 2002; 118: 699–707.
  - 27 Hamann J, Wishaupt JO, Van-Lier RA, Smeets TJ, Breedveld FC, Tak PP. Expression of the activation antigen CD97 and its ligand CD55 in rheumatoid synovial tissue. *Arthritis Rheum* 1999; 42: 650–8.
  - 28 Leemans JC, Te-Valde AA, Florquin S, Bennink RJ, De-Bruin K, Van-Lier RA, *et al*. The epidermal growth factor-seven transmembrane (EGF-TM7) receptor CD97 is required for neutrophil migration and host defense. *J Immunol* 2004; 172: 1125–31.
  - 29 Visser L, De-Vos AF, Hamann J, Melief MJ, Van-Meurs M, Van-Lier RA, *et al*. Expression of the EGF-TM7 receptor CD97 and its ligand CD55 (DAF) in multiple sclerosis. *J Neuroimmunol* 2002; 132: 156–63.
  - 30 Kop EN, Adriaansen J, Smeets TJ, Vervoordelndonk MJ, Van-Lier RA, Hamann J, *et al*. CD97 neutralisation increases resistance to collagen-induced arthritis in mice. *Arthritis Res Ther* 2006; 8: R155.
  - 31 Liu D, Trojanowicz B, Radestock Y, Fu T, Hammje K, Chen L, *et al*. Role of CD97 isoforms in gastric carcinoma. *Int J Oncol* 2010; 36: 1401–8.
  - 32 Hsiao CC, Wang WC, Kuo WL, Chen HY, Chen TC, Hamann J, *et al*. CD97 inhibits cell migration in human fibrosarcoma cells by modulating TIMP-2/MT1- MMP/MMP-2 activity--role of GPS autoproteolysis and functional cooperation between the N- and C-terminal fragments. *FEBS J* 2014; 281: 4878–91.
  - 33 Hsiao CC, Cheng KF, Chen HY, Chou YH, Stacey M, Chang GW, *et al*. Site-specific N-glycosylation regulates the GPS auto-proteolysis of CD97. *FEBS Lett* 2009; 583: 3285–90.
  - 34 Wobus M, Vogel B, Schmucking E, Hamann J, Aust G. N-glycosylation of CD97 within the EGF domains is crucial for epitope accessibility in normal and malignant cells as well as CD55 ligand binding. *Int J Cancer* 2004; 112: 815–22.
  - 35 Lin HH, Chang GW, Davies JQ, Stacey M, Harris J, Gordon S. Autocatalytic cleavage of the EMR2 receptor occurs at a conserved G protein-coupled receptor proteolytic site motif. *J Biol Chem* 2004; 279: 31823–32.
  - 36 Chang GW, Stacey M, Kwakkenbos MJ, Hamann J, Gordon S, Lin HH. Proteolytic cleavage of the EMR2 receptor requires both the extracellular stalk and the GPS motif. *FEBS Lett* 2003; 547: 145–50.
  - 37 Wei W, Hackmann K, Xu H, Germino G, Qian F. Characterization of cis-autoproteolysis of polycystin-1, the product of human polycystic kidney disease 1 gene. *J Biol Chem* 2007; 282: 21729–37.
  - 38 Aricescu AR, Lu W, Jones EY. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr* 2006; 62: 1243–50.
  - 39 Loo T, Patchett ML, Norris GE, Lott JS. Using secretion to solve a solubility problem: high-yield expression in *Escherichia coli* and purification of the bacterial glycoamidase PNGase F. *Protein Expr Purif* 2002; 24: 90–8.
  - 40 Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 2006; 22: 195–201.
  - 41 Wu ST, Zhang Y. LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Res* 2007; 35: 3375–82.
  - 42 Portolano N, Watson PJ, Fairall L, Millard CJ, Milano CP, Song Y, *et al*. Recombinant protein expression for structural biology in HEK 293F suspension cells: a novel and accessible approach. *J Vis Exp* 2014; e51897.
  - 43 Krasnoperov V, Lu Y, Buryanovsky L, Neubert TA, Ichtchenko K, Petrenko AG. Post-translational proteolytic processing of the calcium-independent receptor of alpha-latrotoxin (CIRL), a natural chimera of the cell adhesion protein and the G protein-coupled receptor. Role of the G protein-coupled receptor proteolysis site (GPS) motif. *J Biol Chem* 2002; 277: 46518–26.
  - 44 Davies JQ, Chang GW, Yona S, Gordon S, Stacey M, Lin HH. The role of receptor oligomerization in modulating the expression and function of leukocyte adhesion-G protein-coupled receptors. *J Biol Chem* 2007; 282: 27343–53.
  - 45 Boatright KM, Renatus M, Scott FL, Sperandio S, Shin H, Pedersen IM, *et al*. A unified model for apical caspase activation. *Mol Cell* 2003; 11: 529–41.

- 46 Hsiao CC, Chen HY, Chang GW, Lin HH. GPS autoproteolysis is required for CD97 to up-regulate the expression of N-cadherin that promotes homotypic cell-cell aggregation. *FEBS Lett* 2011; 585: 313–8.
- 47 Chu F. Requirements of cleavage of high mannose oligosaccharides in glycoproteins by peptide N-glycosidase F. *J Biol Chem* 1986; 261: 172–7.
- 48 Zhang K, Zhang J, Gao ZG, Zhang D, Zhu L, Han GW, *et al*. Structure of the human P2Y<sub>12</sub> receptor in complex with an antithrombotic drug. *Nature* 2014; 509: 115–8.
- 49 Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, *et al*. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* 2011; 477: 549–55.
- 50 Zhang D, Gao ZG, Zhang K, Kiselev E, Crane S, Wang J, *et al*. Two disparate ligand-binding sites in the human P2Y<sub>1</sub> receptor. *Nature* 2015; 520: 317–21.