# SCIENTIFIC REPORTS

**OPEN**

# Recombinant production of human $\alpha_2$-macroglobulin variants and interaction studies with recombinant G-related $\alpha_2$-macroglobulin binding protein and latent transforming growth factor-$\beta_2$

Laura Marino-Puertas[1], Laura del Amo-Maestro[1], Marta Taulés[2], F. Xavier Gomis-Rüth[1] & Theodoros Goulas[1]

$\alpha_2$-Macroglobulins ($\alpha_2$Ms) regulate peptidases, hormones and cytokines. Mediated by peptidase cleavage, they transit between native, intact forms and activated, induced forms. $\alpha_2$Ms have been studied over decades using authentic material from primary sources, which was limited by sample heterogeneity and contaminants. Here, we developed high-yield expression systems based on transient transfection in *Drosophila* Schneider 2 and human Expi293F cells, which produced pure human $\alpha_2$M (h$\alpha_2$M) at ~1.0 and ~0.4 mg per liter of cell culture, respectively. In both cases, h$\alpha_2$M was mainly found in the induced form. Shorter h$\alpha_2$M variants encompassing N-/C-terminal parts were also expressed and yielded pure material at ~1.6/~1.3 and ~3.2/~4.6 mg per liter of insect or mammalian cell culture, respectively. We then analyzed the binding of recombinant and authentic h$\alpha_2$M to recombinant latent human transforming growth factor-$\beta_2$ (pro-TGF-$\beta_2$) and bacterial G-related $\alpha_2$M binding protein (GRAB) by surface plasmon resonance, multiple-angle laser light scattering, size-exclusion chromatography, fluorogenic labelling, gel electrophoresis and Western-blot analysis. Two GRAB molecules formed stable complexes of high affinity with native and induced authentic h$\alpha_2$M tetramers. The shorter recombinant h$\alpha_2$M variants interacted after preincubation only. In contrast, pro-TGF-$\beta_2$ did not interact, probably owing to hindrance by the N-terminal latency-associated protein of the cytokine.

$\alpha_2$-Macroglobulins ($\alpha_2$Ms) are large protein inhibitors, which counteract a broad spectrum of endopeptidases. To date, they have been characterized from metazoans and Gram-negative bacteria[1–4]. They are multi-domain molecular traps with comparable structural and biochemical properties, which present related modes of action termed "Venus flytrap" and "snap-trap" mechanisms[5,6]. In both cases, peptidases cut native $\alpha_2$M in a highly flexible bait region, which triggers a massive conformational rearrangement that induces the inhibitor and entraps the peptidase. In some family members, a second event involves a highly reactive $\beta$-cysteinyl-$\gamma$-glutaminyl thioester bond, which is activated by nucleophiles such as lysines and covalently binds the prey peptidase, thus contributing to the stabilization of the enzyme:inhibitor complex. Trapped peptidases are still active but only against small substrates due to steric hindrance[7]. Hence, $\alpha_2$Ms regulate proteolysis in complex biological processes such as

[1]Proteolysis Laboratory, Structural Biology Unit ("Maria de Maeztu" Unit of Excellence), Molecular Biology Institute of Barcelona, Higher Scientific Research Council (CSIC), Barcelona Science Park, Helix Building; Baldiri Reixac, 15-21, 08028, Barcelona, Catalonia, Spain. [2]Scientific and Technological Centers (CCiTUB), University of Barcelona, Lluís Solé i Sabaris, 1-3, 08028, Barcelona, Catalonia, Spain. Correspondence and requests for materials should be addressed to F.X.G.-R. (email: xgrcri@ibmb.csic.es) or T.G. (email: thgcri@ibmb.csic.es)
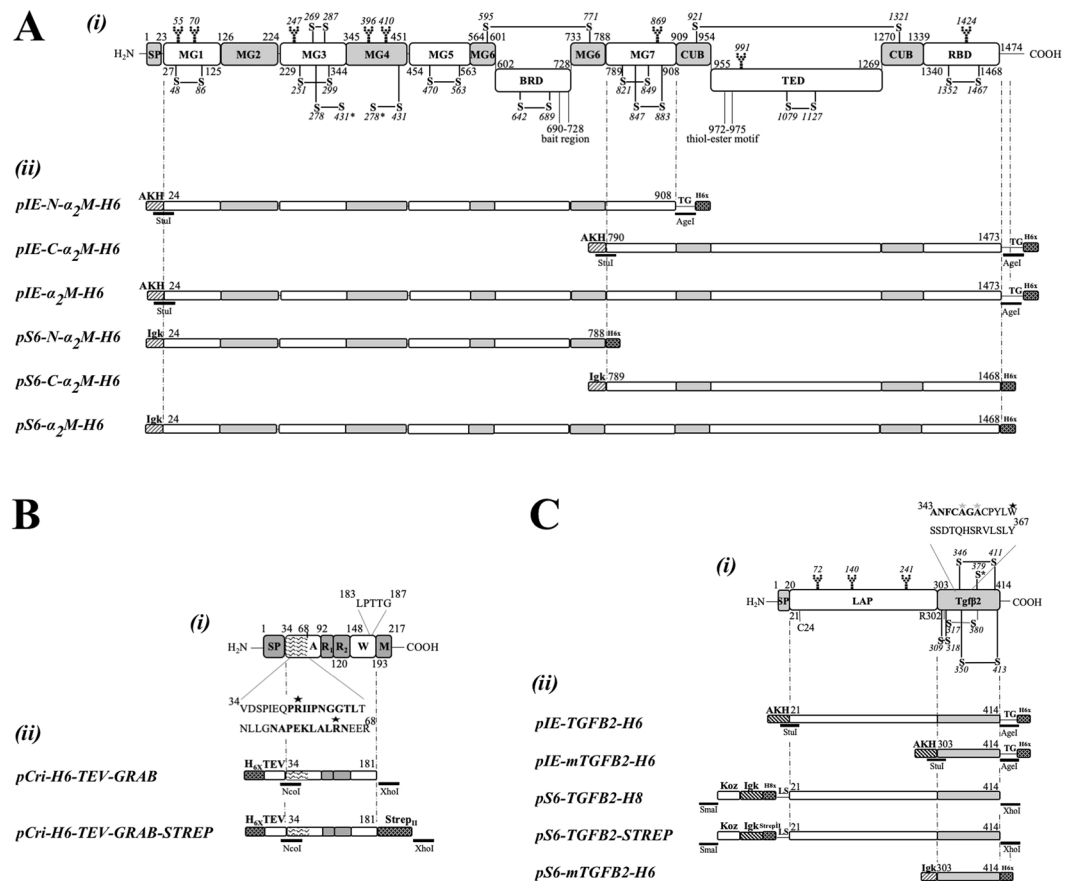
**Figure 1.** Overview of studied proteins. (**A**) Scheme depicting the domain structure of $h\alpha_2M$ (*i*) and the constructs studied (*ii*). The residue numbers correspond to UP P01023. (*i*) Functional regions and domains are the signal peptide (SP); macroglobulin domains 1-to-7 (MG1-MG7); the bait-region domain (BRD); the CUB domain; the thioester domain (TED); and the receptor-binding domain (RBD). Disulfide bonds are shown in black and linked cysteines are labelled. An interchain disulfide is pinpointed with an asterisk and *N*-linked glycosylation sites are highlighted with a sugar chain. (*ii*) $h\alpha_2M$ fusion proteins produced with plasmids pIEx and pCMV-Sport6. The AKH signal peptide sequence, the mouse Ig κ-chain leader sequence, $His_{6\times}$-tags and restriction sites are graphically represented. (**B**) Same as (A) for GRAB (UP Q7DAL7). (*i*) Functional regions and domains are the SP; domain A, with the binding regions of $h\alpha_2M$ hatched and in the inset; repeat regions $R_1$ and $R_2$; the cell-wall attachment site (W), with the cell-wall anchor motif shown in magnification; and the membrane anchor (M). Critical arginine residues for $h\alpha_2M$-binding are indicated by a star ($R^{42}$ and $R^{64}$). (*ii*) GRAB fusion proteins in pCri8a with $His_6$-tag, TEV site and Strep-tag. (**C**) Same as (A) for pro-TGF-$\beta_2$ (UP P61812). (*i*) Functional domains and regions are the SP; the latency associated peptide (LAP); and the mature growth factor moiety (TGF-$\beta_2$). Critical residues in LAP are $C^{24}$, which is involved in the binding of LTBP, and $R^{302}$, required for furin cleavage. Mature TGF-$\beta_2$ segment $A^{343}$-$Y^{367}$ is involved in $h\alpha2M$ binding, important and critical residues are indicated by a grey ($A^{347}$ and $A^{349}$) and a black star ($W^{354}$), respectively. (*ii*) Human pro-TGF-$\beta_2$ fusion proteins produced with plasmids pIEx and pCMV-Sport6. The AKH signal peptide sequence, the Kozac (Koz), the mouse Igκ-chain leader sequence, affinity tags ($His_6$ and Strep) and restriction sites are graphically represented.

digestion, blood homeostasis, signaling, tissue remodeling and defense against toxins and other virulence factors during infection and envenomation[1].

In addition to peptidase binding and inhibition, $\alpha_2Ms$ regulate several other endogenous and exogenous proteins (for a complete list, see[1] and refences therein). Indeed, eukaryotic $\alpha_2Ms$ modify and modulate the activity of cytokines, hormones, growth factors, lipid factors and other proteins, and thus have a great impact on human physiology. A characteristic example is the interaction of human $\alpha_2M$ ($h\alpha_2M$), a 1,474-residue tetrameric multidomain protein (Fig. 1A), with transforming growth factors-$\beta$ (TGF-$\beta$s), a family of ~25-kDa structurally homologous dimeric proteins (Fig. 1C). In mammals, the TGF-$\beta$ family has three members (TGF-$\beta_1$, TGF-$\beta_2$ and TGF-$\beta_3$), which share 70% sequence identity and similar three-dimensional structures[8]. Their biological activity includes growth regulation, transcriptional activation of extracellular-matrix-related genes and chemotactic activity[9,10]. They are primarily regulated by the non-covalently attached N-terminal latency-associated domain (LAP)[11], which acts as a pro-domain in the latent ~100-kDa pro-forms (pro-TGF-$\beta$s). Once in circulation, LAP is removed and TGF-$\beta$ availability is regulated by $h\alpha_2M$, which sequesters most of these cytokines

through a currently unknown mechanism[10,12,13]. What is known is that $h\alpha_2M$ positions $E^{753}$, $E^{737}$ and $D^{742}$ within segment $V^{723}$-$T^{761}$ (numbering according to UniProt [UP] entry P01023) are involved in TGF-$\beta_1$ binding[14] and that induced $h\alpha_2M$ binds the cytokine with higher affinity than the native inhibitor[14].

The functional and structural properties of $h\alpha_2M$ are exploited by pathogens such as *Streptococcus pyogenes* (group A streptococci), which forms stable interactions with $h\alpha_2M$ by a surface protein, the G-related $\alpha_2M$-binding protein (GRAB[15,16]). This 23-kDa protein consists of a Gram-positive membrane anchor motif, a variable number of 28-residue repeats, and a highly-conserved N-terminal domain responsible for the interaction with $h\alpha_2M$ (Fig. 1B). By recruiting native $h\alpha_2M$ to the membrane, GRAB provides *S. pyogenes* with a mechanism to inhibit host peptidases, which protects bacterial surface structures and facilitates progressive dissemination in the infected tissue[15].

These interactions have only been preliminary characterized[17,18], and the mechanisms are still unknown. To shed light on them, we developed eukaryotic expression systems of $h\alpha_2M$ variants and purified the authentic protein from blood. We further used these proteins to study complex formation with GRAB and pro-TGF-$\beta_2$ by several biophysical approaches.

## Materials and Methods

**Construct preparation.** Constructs spanning fragments of the gene coding for $h\alpha_2M$, namely full-length $h\alpha_2M$ and its N- and C-terminal parts (N-$h\alpha_2M$ and C-$h\alpha_2M$; for details on constructs, plasmids, vectors and primers, see Table 1 and Fig. 1), and the coding sequence for GRAB from *Streptococcus pyogenes* serotype M1 (UP Q7DAL7) were amplified with primers that introduced either restriction sites for directional cloning or overhangs for restriction-free cloning. The vectors used were pCri-8a[19] for bacterial expression, pIEx (Novagen) for expression in *Drosophila melanogaster* Schneider 2 embryonic cells (S2; Gibco), and pCMV-Sport 6 (Thermo Scientific) for expression in human Expi293F™ cells (Gibco). Polymerase chain reaction (PCR) primers and DNA modifying enzymes were purchased from Sigma-Aldrich and Thermo Scientific, respectively. PCR was performed using Phusion High Fidelity DNA polymerase (Thermo Scientific) according to the manufacturer's instructions and following a standard optimization step by thermal gradient in each reaction. Mutants were generated by a modified version of the previously described procedure[20]. DNA was purified with the OMEGA Biotek Purification Kit according to the manufacturer's instructions, and all constructs were verified by DNA sequencing.

**Cell-culture media.** S2 and Expi293F cells were adapted to grow in suspension in Sf-900™ II SFM culture medium (Gibco) and FreeStyle™ F17 expression medium (Gibco) with 0.2% Pluronic F-68 (Gibco) *plus* 8 mM L-glutamine (Gibco), respectively. Both growth media were supplemented with 0.5 µg/mL of the antimycotic Fungizone, 100 units/mL of penicillin, and 100 µg/mL of streptomycin sulfate (Gibco).

**Cell-culture growth.** S2 cells were cultivated in TubeSpin bioreactor tubes (TS50 for 5-to-10-mL cultures and TS600 for 100-to-200-mL cultures; Techno Plastic Products AG) as previously described[21]. Cells were passaged three times per week to a final density of $4 \times 10^6$ cells/mL. The cultures were incubated at 28 °C in a shaker (Brunswick Scientific Innova) under agitation at 220 rpm.

Expi293F cells were cultivated in 125-mL or 1000-mL polycarbonate Erlenmeyer flasks (FPC0125S and FPC1000S, respectively; Tri Forest Labware) for 25-to-30-mL and 100-to-250-mL cultures, respectively. Cells were subcultured three times per week to a final density of $0.3–0.5 \times 10^6$ cells/mL and kept in suspension at 150 rpm in a Multitron Cell Shaker Incubator (Infors HT) at 37 °C in a modified atmosphere (8% $CO_2$ and 85% of relative humidity). Cell densities and viability were determined by the trypan blue exclusion test[22].

**Cell-culture transfection.** Linear 25-kDa polyethylenimine (PEI; Polysciences Europe GmbH) was prepared in Milli-Q water at a concentration of 1 mg/mL and pH 7.0. The solution was filter-sterilized and stored at −20 °C. Plasmid DNA was produced in *Escherichia coli* DH5α cells, purified with the GeneJET Plasmid Maxiprep Kit (Thermo Scientific), and stored at −20 °C in sterile Milli-Q water at 1 mg/mL.

For transfection, S2 cells were centrifuged and resuspended in prewarmed fresh medium to a cell density of $15 \times 10^6$ cells/mL. A mixture of 0.6 µg DNA (see Fig. 1 and Tables 1) and 2 µg PEI per $1 \times 10^6$ cells and per prewarmed transfection volume was pre-incubated for 15–20 min at room temperature and then added dropwise to the cell cultures. These were further incubated for 1 hour at 28 °C and 220 rpm, subsequently diluted with prewarmed fresh medium to $5 \times 10^6$ cells/ml and harvested after seven days for protein purification.

For mammalian cultures, Expi293F cells were transfected at a cell density of $1 \times 10^6$ cells/mL with a mixture of 1 mg of DNA (see Table 1) and 3 mg of PEI in 20 mL of Opti-MEM Medium (Gibco) per liter of expression medium. The DNA-PEI mixture was incubated at room temperature for 15–20 min and then added dropwise to the cell cultures, which were harvested after three days for protein purification.

**Bacterial expression.** For the recombinant overexpression of N-terminally hexa-histidine (His$_6$)-tagged GRAB with a tobacco-etch virus peptidase (TEV) recognition sequence, with or without an additional C-terminal Streptactin ®II tag (Strep-tag; IBA Life Sciences), plasmid pCRI8a[19] was transformed into *E. coli* BL21 (DE3) cells (Novagen[23]), and cultures were grown in lysogeny broth supplemented with 30 µg/mL kanamycin. After initial growth at 37 °C to an $OD_{600} \approx 0.6$, cultures were cooled to 20 °C, and protein expression was induced with 0.4 mM isopropyl-β-D-thiogalactopyranoside for 18–20 hours.

**Protein purification.** Protein purification steps were performed at 4 °C if not otherwise stated. For GRAB purification, bacterial cells were collected by centrifugation at $6,000 \times g$ for 30 min, washed in buffer A (50 mM Tris-HCl, 250 mM sodium chloride, pH 7.5) and resuspended in the same buffer *plus* 20 mM imidazole, Complete EDTA-free Peptidase Inhibitor Cocktail Tablets and DNase I (both from Roche Diagnostics). Cells were lyzed with a cell disrupter (Constant Systems) at a pressure of 1.35 kbar, cell debris was removed by centrifugation at

| Plasmid name | Protein | Parental DNA | Forward -primer* | Reverse-primer* | Protein sequence** | Tags*** | Comments |
|---|---|---|---|---|---|---|---|
| pIE-h$\alpha_2$M-H6 | h$\alpha_2$M | Human c-DNA pIEx vector | CATTAGGCCTC AGTCTCTGGAAA ACCGCAGTATATG | CATTACCGGTA GGATTTCC AAGATCTTTG | $S^{24}$-$N^{1473}$ + **PTG** + **H$_{6\times}$** | C-t H$_{6\times}$ | Full-length h$\alpha_2$M in S2 cells. The gene was inserted by directional cloning (between *Stu*I and *Age*I) into the pIEx vector in frame with the AKH signal peptide sequence. |
| pIE-N-h$\alpha_2$M-H6 | N-h$\alpha_2$M | pIE-h$\alpha_2$M-H6 | GATCTTGGAAAT CCTACCGGTCAT CATCAC | CAATACCGGT TTCAGGTTC AACCAACAGAG | $S^{24}$-$E^{908}$ + **TG** + **H$_{6\times}$** | C-t H$_{6\times}$ | As above for the N-terminal half of h$\alpha_2$M. |
| pIE-C-h$\alpha_2$M-H6 | C-h$\alpha_2$M | pIE-h$\alpha_2$M-H6 | CAATAGGCCTCA CAGCCCTTCTT TGTGGAGCTC | CAATAGGCCT CAGCGATGA TGACGAAAG | $Q^{790}$- $N^{1473}$ + **TG** + **H$_{6\times}$** | C-t H$_{6\times}$ | As above for the C-terminal half of h$\alpha_2$M. |
| pS6-h$\alpha_2$M-H6 | h$\alpha_2$M | pIE-h$\alpha_2$M-H6 pCMV-Sport 6 vector | TGGGTTCCAGGTT CCACTGGTGACTCAGTC TCTGGAAAACCGCAGTAT | CGCCTAATGGTGAT GGTGATGGTGGCTGCA AGGAGCATTGTACTCAGC | $S^{24}$-$S^{1468}$ + **H$_{6\times}$** | C-t H$_{6\times}$ | Full-length h$\alpha_2$M in Expi293F cells. The gene was inserted by restriction-free cloning into the pCMV-Sport 6 vector in frame with the Ig $\kappa$ leader sequence. |
| pS6-N-h$\alpha_2$M-H6 | N-h$\alpha_2$M | pIE-N-h$\alpha_2$M-H6 pCMV-Sport 6 vector | TGGGTTCCAGGTT CCACTGGTGACTCAGTC TCTGGAAAACCGCAGTAT | CGCCTAATGGT GATGGTGATGGTGGG CTCGGAGAGA GGCAGTGGAAGA | $S^{24}$-$A^{788}$ + **H$_{6\times}$** | C-t H$_{6\times}$ | As above for the N-terminal half of h$\alpha_2$M. |
| pS6-C-h$\alpha_2$M-H6 | C-h$\alpha_2$M | pIE-h$\alpha_2$M-H6 pCMV-Sport 6 vector | TGGGTTCCAGGTTCCACT GGTGACTTCCAGC CCTTCTTTGTGGAGCTC | CGCCTAATGGTGAT GGTGATGGTGG CTGCAAGGAGC ATTGTACTCAGC | $F^{789}$-$S^{1468}$ + **H$_{6\times}$** | C-t H$_{6\times}$ | As above for the C-terminal half of h$\alpha_2$M. |
| pCri-H6-TEV-GRAB | GRAB | Synthetic DNA pCri8a vector | CAATCCATGGTT GATAGCCCGATTG | CAATCTCGAGT TAATTAACGT TCTGACGTT | **GAM** + $V^{34}$-$N^{181}$ | N-t H$_{6\times}$ + TEV | Synthetic gene of GRAB optimized for expression in *Escherichia coli* inserted into the pCri8a vector[19] by directional cloning between the *Nco*I and *Xho*I restriction sites. |
| pCri-H6-TEV-GRAB-STREP | GRAB | pCri-H6-TEV-GRAB | ATGCCCATGGT TGATAGCCCG | CGAATTGTGGATGGCTC CAACCTCCATTAACGT TCTGACGTTC; CTTCCACCTCCAGA ACCTCCACCCTTTC GAATTGTGGATGGCTCC; GTGGATGGCTCCATGCG CTACCTCCACTTCCACCT CCAGAACC; GCATCTCGAG TTACTTTT CGAATTGT GGATGGCTC CATGCGC | **GAM** + $V^{34}$-$N^{181}$ + **G GWSHPQFEKGGG SGGGSGGSAWS HPQFEK** | N-t H$_{6\times}$ + TEV-(protein)-Strep | This construct was obtained from pCri-H6-TEV-GRAB by four consecutive PCR reactions to introduce a C-terminal Strep-tag. |
| pIE-TGFB2-H6 | pro-TGF-$\beta_2$ | Human c-DNA pIEx vector | CAATAGGCCTT GTCTACCTG CAGCACACTC | CAATACCGGTGC TGCATTTGCAAG ACTTTAC | $L^{21}$-$S^{414}$ + **TG** + **H$_{6\times}$** | C-t H$_{6\times}$ | Pro-TGF-$\beta_2$ in S2 cells. The gene was inserted by directional cloning (between *Stu*I and *Age*I) into the pIEx vector in frame with AKH signal peptide sequence |
| pIE-mTGFB2-H6 | TGF-$\beta_2$ | Human c-DNA pIEx vector | CAATAGGCCTC AGTTTGGAT GCGGCCTATTG | CAATACCGGTG CTGCATTTGCA AGACTTTAC | $A^{303}$-$S^{414}$ + **TG** + **H$_{6\times}$** | C-t H$_{6\times}$ | As above for mature TGF-$\beta_2$. |
| pS6-TGFB2-H8 | pro-TGF-$\beta_2$ | pIE-TGFB2-H6 pCMV-Sport 6 vector | TCACCACCACCATCATCTCA GCCTGTCTACCTGCAGCA; GGTTCCACTGGTGACCACC ACCATCACCACCACCATC; GGGTACTGCTGCTCTGGGTT CCAGGTTCCACTGGTGAC; GACAGACACACTCCTGCT ATGGGTACTGCTGCTC; CAATCCCGGGGCCACCAT GGAGACAGACACACTCC | CAATCTCGAGCT AGCTGCATTTGC AAGACTTTAC | **H$_{8\times}$** + **LS** + $L^{21}$-$S^{414}$ | N-t H$_{8\times}$ | Pro-TGF-$\beta_2$ in Expi293F cells. See[8] for details. |
| pS6-TGFB2-STREP | pro-TGF-$\beta_2$ | pS6-TGFB2-H8 | GGTGGAGG TTCTGGAG GTGGAAGT GGAGGTAGCG CATGGAGCC ATCCACAA TTCGAA AAGCTCA GCCTGTC TACCTGC | CTTTTCGAATTGTG GATGGCTCCAGTCA CCAGTGGAACCTGGA ACCCAGAGCAG | **WSHPQFEK GGGSGGGS GGSAWSHPQ FEKLS** + $L^{21}$-$S^{414}$ | N-t Strep | Pro-TGF-$\beta_2$ in Expi293F cells. The parental plasmid was modified by opposite primers to replace the N-terminal histidine-tag with a Strep-tag. See[8] for details. |
| pS6-mTGFB2-H6 | TGF-$\beta_2$ | pIE-mTGFB2-H6 pCMV-Sport 6 vector | GTTCCAGGTTCC ACTGGTGACG CTTTGGATG CGGCCTATTGC | CCTAATGGT GATGGTGAT GGTGGCTGC ATTTGCAAGA CTTTACA | $A^{303}$-$S^{414}$ + **H$_{6\times}$** | C-t H$_{6\times}$ | Mature TGF-$\beta_2$ in Expi293F cells. The coding gene extracted from the parental plasmid was inserted into the pCMV-Sport 6 vector by restriction-free cloning between the Ig $\kappa$ leader sequence and the C-terminal histidine-tag. |

**Table 1.** Constructs, primers, plasmids and proteins. All constructs are for extracellular expression of the respective proteins. *Restriction-site sequences and overhangs for restriction-free cloning are underlined. **Peptide sequence of the expressed protein after fusion-tag removal. Amino acids derived from the construct are in bold. See also Fig. 1. ***Fused tags at the carboxy-terminus (C-t) or the amino-terminus (N-t). AKH, adipokinetic hormone; TEV, tobacco-etch virus peptidase; Ig $\kappa$, immunoglobulin $\kappa$.
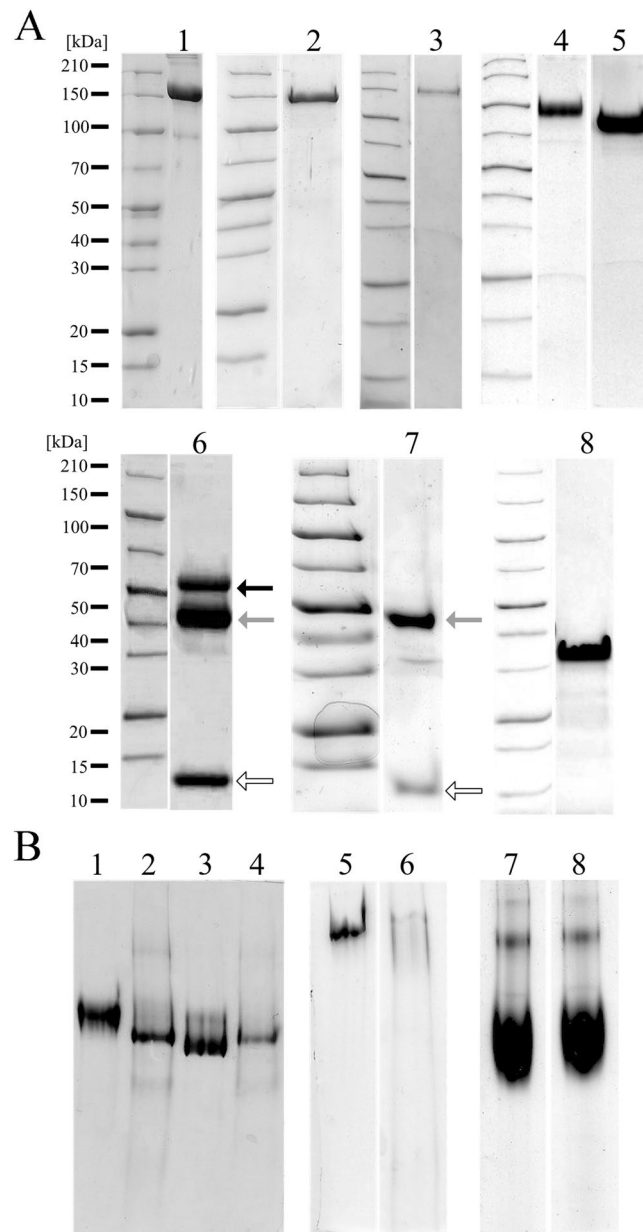
**Figure 2.** Recombinant protein production and purification. (**A**) SDS-PAGE analysis of wild-type and recombinant proteins. Lanes: 1, native authentic hα$_2$M; 2, recombinant hα$_2$M from S2 cells; 3, recombinant hα$_2$M from Expi293F cells; 4, N-terminal half of hα$_2$M (N-hα$_2$M); 5, C-terminal half of hα$_2$M (C-hα$_2$M); 6, pro-TGF-β$_2$ produced in Expi293F cells according to[8]. Arrows indicate pro-TGF-β$_2$ (black), LAP (grey) and mature TGF-β$_2$ (white); 7, pro-TGF-β$_2$ digested by furin; 8, GRAB. (**B**) Native-PAGE analysis of wild-type and recombinant proteins. Lanes: 1 and 3, native and methylamine-induced authentic hα$_2$M; 2 and 4, native and methylamine-induced recombinant hα$_2$M from S2 cells; 5 and 6, native authentic hα$_2$M and recombinant hα$_2$M from Expi293F cells; 7 and 8, native and induced recombinant C-α$_2$M expressed from Expi293F cells.

30,000 × *g* for 1 hour, and the supernatant containing GRAB was kept for subsequent purification steps. For the hα$_2$M variants produced in S2 and Expi293F systems, cells were removed by centrifugation at 2,800 × *g* for 20 min and the supernatant was used for subsequent purification steps.

Supernatants containing the proteins of interest were incubated for 20 min (expression in insect cells) or 1 hour (expression in mammalian cells) with nickel-nitrilotriacetic acid resin (Ni-NTA; Invitrogen), which was subsequently loaded onto an open column for batch purification (Bio-Rad), washed extensively with buffer A *plus* 20 mM imidazole, and eluted with buffer A *plus* 300 mM imidazole (direct Ni-NTA). For GRAB, eluted samples were then dialyzed overnight against buffer A *plus* 1 mM 1,4-dithio-DL-threitol (DTT) in the presence of His$_6$-tagged TEV at a peptidase:protein weight ratio of 1:100 and 1 mM DTT. The resulting cleavage left additional residues (glycine-alanine-methionine) at the N-terminus of the target proteins due to the cloning strategy (see Table 1). Digested samples were passed several times through Ni-NTA resin previously equilibrated with buffer A *plus* 20 mM imidazole to remove His$_6$-tagged molecules and the flow-through containing untagged GRAB was collected (reverse Ni-NTA).

| Protein sample | Molecular mass (kDa) |
|---|---|
| Native h$\alpha_2$M | 680.6 ± 1.8 |
| Native h$\alpha_2$M + GRAB | 707.8 ± 3.4 |
| Induced h$\alpha_2$M | 684.1 ± 2.7 |
| Induced h$\alpha_2$M + GRAB | 710.6 ± 1.5 |
| GRAB | 15.5 ± 0.0 |
| pro-TGF-$\beta_2$ | 105.4 ± 0.6 |

**Table 2.** Molecular masses determined by SEC-MALLS. Values are represented as means and standard deviations of three replicates.

In all cases, proteins eluted from direct and reverse Ni-NTA chromatographies were dialyzed overnight against buffer B (20 mM Tris-HCl, 5 mM sodium chloride, pH 7.5) and further purified by ionic-exchange chromatography (IEC) on a TSKgel DEAE-2SW column (TOSOH Bioscience) equilibrated with buffer B. A gradient of 2–30% buffer C (20 mM Tris-HCl, 1 M sodium chloride, pH 7.5) was applied over 30 mL, and samples were collected and pooled. Subsequently, each pool was concentrated by ultrafiltration and subjected to size-exclusion chromatography (SEC) in Superdex 75 10/300 (GRAB and pro-TGF-$\beta_2$), Superdex 200 10/300 (N-h$\alpha_2$M and C-h$\alpha_2$M) or Superose 6 10/300 (full-length recombinant h$\alpha_2$M) columns (GE Healthcare Life Sciences) in buffer D (20 mM Tris-HCl, 150 mM sodium chloride, pH 7.5). Strep-tagged GRAB was purified by affinity chromatography with Streptactin®XT Superflow Suspension resin (IBA Life Sciences) and eluted with buffer E (100 mM Tris·HCl, 150 mM sodium chloride, pH 8.0) at a further 50 mM in biotin. IEC and SEC purification steps followed as above.

Authentic full-length h$\alpha_2$M was isolated from blood plasma from individual donors and purified essentially as described previously[17,24,25]. Briefly, plasma was subjected to sequential precipitation steps with 4–12% PEG 4,000, and the final precipitate containing h$\alpha_2$M was reconstituted in 20 mM sodium phosphate at pH 6.4. Partially purified h$\alpha_2$M was captured with a zinc-chelating resin (G-Biosciences), washed with buffer F (50 mM sodium phosphate, 250 mM sodium chloride, pH 7.2) *plus* 10 mM imidazole and eluted in the same buffer *plus* 250 mM imidazole and 100 mM EDTA. The protein was first passed through a PD10 desalting column (GE Healthcare Life Sciences) previously equilibrated with 20 mM HEPES, pH 7.5 and then subjected to an IEC step in a Q Sepharose column (2.5 × 10 cm; GE Healthcare Life Sciences), previously equilibrated with 15% buffer G (20 mM HEPES, 1 M sodium chloride, pH 7.5). A gradient of 20–30% buffer G was applied for 150 min and fractions were collected. Collected samples were dialyzed overnight against buffer H (20 mM sodium phosphate, 5 mM sodium chloride, pH 7.4) and further purified by IEC in a TSKgel DEAE-2SW column, previously equilibrated with buffer H. A gradient of 7–20% buffer I (20 mM sodium phosphate, 1 M sodium chloride, pH 7.4) was applied over 30 mL, and samples were collected and pooled. Subsequently, each pool was concentrated and subjected to a final polishing step by SEC in a Superose 6 10/300 column in buffer J (20 mM sodium phosphate, 150 mM sodium chloride, pH 7.4).

Protein identity and purity were assessed by 10–15% Tricine sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE[26]) stained with Coomassie Brilliant Blue, peptide mass fingerprinting of tryptic protein digests, N-terminal sequencing through Edman degradation, and mass spectrometry. The latter three were carried out at the Protein Chemistry Service and the Proteomics Facilities of the Centro de Investigaciones Biológicas (Madrid, Spain). Ultrafiltration steps were performed with Vivaspin 15 and Vivaspin 500 filter devices of 10- to 50-kDa cut-off (Sartorius Stedim Biotech). Protein concentrations were estimated by measuring the absorbance at 280 nm in a spectrophotometer (NanoDrop) and applying the respective theoretical extinction coefficients. Concentrations were also measured by the BCA Protein Assay Kit (Thermo Scientific) with bovine serum albumin fraction V (BSA; Sigma-Aldrich) as a standard. Induced h$\alpha_2$M was obtained by treating native h$\alpha_2$M in buffer D with 200 mM methylamine hydrochloride for one hour at room temperature. Subsequently, the sample was dialyzed against the same buffer D.

Human pro-TGF-$\beta_2$ (UP P61812) constructs (see Table 1 and Fig. 1C) were produced in S2 and Expi293F cells and purified as reported elsewhere[8]. Production of mature TGF-$\beta_2$ with a C-terminal His$_6$-tag (see Table 1) was assayed with the insect and human systems, which included harvesting periods of seven and three days, respectively. Supernatants were collected after the centrifugation at 2,800 × *g* for 20 min and the purification steps were, first a direct Ni-NTA in buffer A *plus* 20 mM imidazole for the wash step, and *plus* 300 mM imidazole for the elution; and finally purified by SEC with a Superdex 75 10/300 column in buffer D.

**Protein labeling.** GRAB and pro-TGF-$\beta_2$ were labelled with fluorogenic sulfosuccinimidyl-7-amino-4-methylcoumarin-3-acetate (Sulfo-NHS-AMCA; Thermo Scientific) according to the manufacturer's instructions with a 10–15 molar excess of reagent over protein in buffer J for 1 hour at room temperature. Thereafter, the proteins were extensively dialyzed against buffer J to remove non-reacted dye. To assess binding, labelled GRAB or pro-TGF-$\beta_2$ were mixed with authentic h$\alpha_2$M (native and induced) or recombinant fragments N-h$\alpha_2$M and C-h$\alpha_2$M at a 4:1 molar ratio, incubated in buffer J for two hours at 37 °C, and analyzed by 10% native PAGE[27]. Gel fluorescence was visualized in a gel reader (G:BOX F3 Gel Doc System, Syngene) and the fluorescence was measured ($\lambda_{ex}$ = 345–350 nm and $\lambda_{em}$ = 440–460 nm). Negative controls (unlabeled proteins) were included in each experiment. After fluorescence detection, native gels were stained with Coomassie Brilliant Blue (Thermo Scientific) to detect the negative controls.

**Multi-angle laser light scattering.** Multi-angle laser light scattering in a Dawn Helios II apparatus (Wyatt Technologies) coupled to a SEC Superose 6 10/300 column (SEC-MALLS) equilibrated in buffer J at 25 °C was performed at the joint IBMB/IRB Crystallography Platform, Barcelona Science Park (Catalonia, Spain) to analyze binding of GRAB or pro-TGF-$\beta_2$ to native or induced authentic h$\alpha_2$M at a molar ration of 4:1. ASTRA 7 software
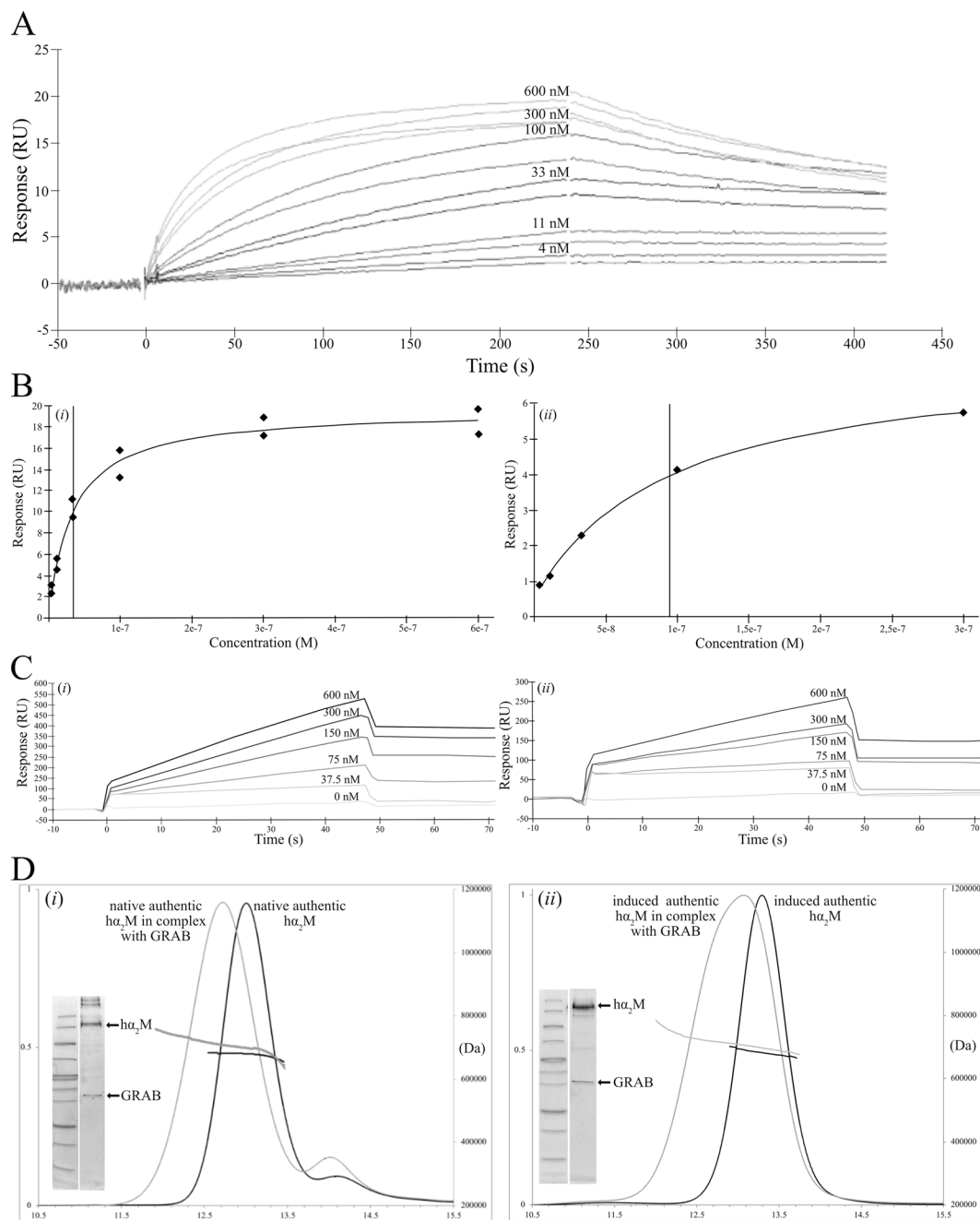
**Figure 3.** Interaction of GRAB and pro-TGF-$\beta_2$ with h$\alpha_2$M variants. (**A,B**) Surface-plasmon resonance sensorgrams of the interaction of native or induced authentic h$\alpha_2$M with GRAB. Multi-cycle run for native h$\alpha_2$M with GRAB (A) and corresponding plot of the steady-state response (B, *i* and *ii*, for native and induced h$\alpha_2$M, respectively). Different h$\alpha_2$M concentrations were assayed to determine the rate constants that describe the kinetics and the equilibrium constants for complex strength (see also Tables 3 and 4). The vertical line in the plots of steady-state response indicates the value of the calculated equilibrium dissociation constant $K_D$. (**C**) Sensorgrams of the interaction of N-h$\alpha_2$M (*i*) and C-h$\alpha_2$M (*ii*) with GRAB. Proteins were premixed, incubated at 37 °C for 1 h, injected over the chip, and the response was measured. (**D**) SEC-MALLS analysis of complex formation between GRAB and native (*left*) and induced (*right*) authentic h$\alpha_2$M showing the measured molecular mass distribution. Inserted figures within graphs show the SDS-PAGE analysis of the respective purified complexes.

---

(Wyatt Technologies) was used for data processing and analysis, for which a dn/dc value typical for proteins (0.185 mL/g) was assumed. All experiments were performed in triplicate.

**Western blot analyses.** Protein samples were separated by 10% SDS-PAGE, transferred to Hybond ECL nitrocellulose membranes (GE Healthcare Life Sciences), and blocked for two hours under gentle stirring at room temperature with 50 mL of blocking solution (phosphate buffered saline; PBS) *plus* 0.1% Tween 20 and 5%

| Protein sample | $k_a$ (M$^{-1}$ s$^{-1}$) | $k_d$ (s$^{-1}$) | $K_D$ (M) | $R_{max}$ (RU) | $\chi^2$ |
|---|---|---|---|---|---|
| Native h$\alpha_2$M + GRAB | $1.32 \times 10^{+5}$ | $1.90 \times 10^{-3}$ | $1.43 \times 10^{-8}$ | 18.51 | 1.23 |

**Table 3.** Kinetic rates and equilibrium constants of the interaction between native authentic h$\alpha_2$M and GRAB. Constants were calculated from the corresponding plot assuming a 1:1 interaction model (two GRAB molecules per h$\alpha_2$M dimer), see Fig. 3A; $k_a$, association rate constant; $k_d$, dissociation rate constant; $K_D$, equilibrium dissociation constant.

| Protein sample | $K_D$ (M) | $R_{max}$ (RU) | $\chi^2$ |
|---|---|---|---|
| Native h$\alpha_2$M + GRAB | $3.45 \times 10^{-8}$ | 18.94 | 1.17 |
| Induced h$\alpha_2$M + GRAB | $9.46 \times 10^{-8}$ | 6.82 | 0.01 |

**Table 4.** Equilibrium constants of the interaction between native or induced authentic h$\alpha_2$M and GRAB. Values were derived from the corresponding plot of steady state response against concentration assuming a 1:1 model (one GRAB molecule per h$\alpha_2$M dimer), see Fig. 3B.
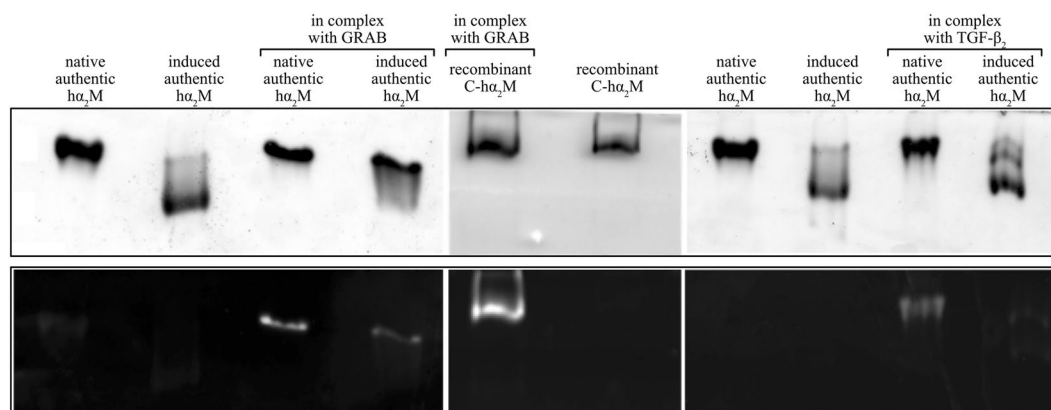


**Figure 4.** Analysis of complex formation between h$\alpha_2$M variants and GRAB or pro-TGF-$\beta_2$. Complexes were separated by native-PAGE. GRAB or TGF-$\beta_2$ labelled with fluorogenic Sulfo-NHS-AMCA were visualized in a gel reader (*lower panels*) and then stained with Coomassie Brilliant Blue (*upper panels*).

BSA. His$_6$-tagged proteins were detected by immunoblot analysis using the monoclonal His-HRP Conjugated Antibody (Santa Cruz Biotechnology) diluted 1:5,000 in PBS *plus* 0.1% Tween 20. Strep-tagged proteins were detected with the Streptavidin-Peroxidase Conjugated Antibody from *Streptomyces avidinii* (Sigma-Aldrich) diluted 1:1,000 in PBS *plus* 0.1% Tween 20 and 1% BSA. Complexes were detected using an enhanced chemiluminescence system (Super Signal West Pico Chemiluminescent; Pierce) according to the manufacturer's instructions. Membranes were exposed to Hyperfilm ECL films (GE Healthcare Life Sciences).

**Proteolytic inhibition assays.** Inhibition assays against protein substrates were performed in a microplate fluorimeter (Infinite M200, TECAN) in 200 μL reaction volumes with the fluorescence-based EnzCheck Assay Kit containing BODIPY FL-casein ($\lambda_{ex} = 505$ nm and $\lambda_{em} = 513$ nm) as fluorescein conjugate (Invitrogen) at 10 μg/mL in buffer D. Inhibition was measured after preincubation of a two-fold molar excess of authentic or recombinant h$\alpha_2$M with trypsin (0.25 μg) for 15 min at room temperature. The substrate was added to the reaction mixture and the residual tryptic activity was measured over a period of two hours.

**Thiol quantification.** Detection of free sulfhydryl groups was performed with the Fluorometric Thiol Assay Kit (ab112158 assay; Abcam) following the manufacturer's instructions and using glutathione as a standard for the dose response curve. The fluorescent signal was measured in a microplate fluorimeter (Infinite M200, TECAN) at $\lambda_{ex} = 490$ nm and $\lambda_{em} = 520$ nm in 96-well plates containing 100 μL reaction volumes (50 μL of assay reaction mixture *plus* 50 μL of glutathione-standard or test samples) in duplicate. Fluorescence was measured after preincubation of authentic h$\alpha_2$M (0.39 μM) or C-h$\alpha_2$M obtained from human cells (1.6 μM), with or without treatment with methylamine for 10, 20, 30, 45 and 60 min, at room temperature.

**Surface plasmon resonance and kinetic data analysis.** The binding kinetics (association and dissociation) and affinity (complex formation at the equilibrium) of GRAB or pro-TGF-$\beta_2$ (ligands) with native authentic h$\alpha_2$M, induced authentic h$\alpha_2$M, recombinant N-h$\alpha_2$M or recombinant C-h$\alpha_2$M (analytes) were studied by surface plasmon resonance with a Biacore$^{TM}$ T200 Biosensor System (GE Healthcare Life Sciences) at the Scientific and Technological Centers of the University of Barcelona (Catalonia, Spain). To bind ligands provided with a Strep-tag, Streptactin®XT (IBA LifeSciences) was immobilized at 25 °C on the surface of the four flow cells of a sensor chip CM5 series S (GE Healthcare Life Sciences) at 3,000 response units (RU) through amine coupling, as

described previously[22]. Subsequently, Strep-tagged GRAB (at 9.7 nM) or pro-TGF-$\beta_2$ (at 19.0 nM) in HBNS buffer (10 mM HEPES, 150 mM sodium chloride, pH 7.4) were immobilized at low RU density on different flow cells of the chip by virtue of the strong interaction between the Strep-tag and streptactin at 5 μL/min for 24 sec at 37 °C. To monitor association, the immobilized ligands were then exposed to the analytes at different concentrations in HBNS (4–600 nM for native and induced authentic h$\alpha_2$M; 75–1,200 nM for N-h$\alpha_2$M and C-h$\alpha_2$M), which were injected at 30 μL/min for 120–240 sec at 37 °C. Thereafter, HBNS was injected for analyte dissociation from the immobilized ligands for 90–300 sec. To dissociate bound ligands and regenerate the chip surface, 3 M guanidine hydrochloride was injected at 30 μL/min for 30 sec after each cycle. These experiments were double referenced by keeping the first flow cell without ligand, and by an injection step at analyte concentration zero. The affinity analysis was performed by plotting binding responses in the steady-state region of the sensorgrams ($R_{eq}$) against analyte concentrations to determine the overall equilibrium dissociation constant ($K_D$). Sensorgrams were analyzed with the BIAEVALUATION program v. 3.0 (GE Healthcare Life Sciences) and fitted to a 1:1 Langmuir interaction model. The likelihood of fitting was assessed through the $\chi^2$ statistical parameter[28].

In a separate qualitative experiment, ligands GRAB (at 120 nM) and pro-TGF-$\beta_2$ (at 950 nM) were premixed with the analytes at different concentrations (2–150 nM for native and induced authentic h$\alpha_2$M; 38–600 nM for N-h$\alpha_2$M and C-h$\alpha_2$M) and incubated for one hour at 37 °C. Subsequently, the mixtures were injected at 15 μL/min at 37 °C according to a published multicycle method[29]. The binding was measured through the increase in RU after injection of the premixes and the stability of the resultant complexes through their elution with buffer HBNS at a flow rate of 30 μL/min. Ligand solutions without analyte were used as negative controls of complex formation and the sensor surface was regenerated after each sample injection.

## Results and Discussion

### Biochemical characterization of the recombinant proteins.

Authentic h$\alpha_2$M has been routinely isolated from blood serum, where it is found at an excess of 2–4 mg/mL but is rather heterogeneous as to conformational state, glycosylation and presence of contaminants[17,18]. Native recombinant h$\alpha_2$M was obtained from immortalized myelogenous leukemia cell line K-562 but the yield was not reported[14]. Therefore, efforts were made here to develop a system for heterologous expression of the protein with high yield, purity and homogeneity, as well as the necessary flexibility to engineer the protein at will. Full-length h$\alpha_2$M with a C-terminal His$_6$-tag was expressed in S2 insect cells using a standard transfection protocol[30] and the signal peptide of the adipokinetic hormone (AKH) for secretion to the extracellular environment. After seven days of expression and harvesting of the supernatant, the protein was purified by affinity chromatography, IEC and SEC steps with yields of up to ~1.0 mg of pure protein per liter of expression medium (Fig. 2A). The protein migrated as a tetramer of ~690 kDa according to SEC (data not shown). Its electrophoretic mobility in native-PAGE was similar to that of induced authentic h$\alpha_2$M (Fig. 2B), which migrates faster than the native protein[31]. Chemical treatment with methylamine, which mimics the transition from native to induced h$\alpha_2$M by opening the reactive thioester bond to produce a free cysteine[31], did not have any effect on protein mobility. Consistently, the protein could not inhibit trypsin activity against a fluorogenic protein substrate, even at 10-fold molar excess. We conclude that recombinant h$\alpha_2$M produced in insect cells was in the induced form, which does not permit the physiological entrance and entrapment of attacking peptidases[4], similarly to a previous report of a baculovirus expression system[32]. Moreover, the thioester bond was either not formed or it was opened after secretion into the extracellular environment by nucleophiles from the expression medium. Unfortunately, we could not evaluate this possibility as the composition of the commercial medium that was used is not available. However, the latter hypothesis seems more plausible given that it is reported that insects can produce thioester-containing proteins[33].

We next developed a transient expression system based on Expi293F cells, which derived from the HEK293 human embryonic kidney cell line and were cultured and harvested at 37 °C for three days. The protein was furnished with the leader sequence of mouse immunoglobulin κ (Ig κ) for secretion and produced ~0.4 mg of pure protein per liter of expression medium (Fig. 2A). The protein migrated as a tetramer in SEC and showed electrophoretic mobility in native-PAGE between native and induced authentic h$\alpha_2$M (Fig. 2B). Consistently, its capacity to inhibit trypsin was 35% of native authentic h$\alpha_2$M. Together, these data indicate that the recombinant protein is partly native but mainly induced. Previous studies had indicated that thioester formation is a spontaneous process triggered by the packing energy of the polypeptide chain during folding in mammals[34]. Therefore, the limited ability of the Expi293F system to produce native protein was attributed, as in the insect cell system above, to the expression medium rather than to the lack of crucial cell machinery for proper thioester bond formation.

Then we expressed shorter variants of h$\alpha_2$M in the insect and mammalian systems (Fig. 2A). N-h$\alpha_2$M spanned from macroglobulin-like (MG) domain 1 (MG1) to MG7 in the insect cell system and from MG1 to MG6 in the mammalian system. C-h$\alpha_2$M ranged from MG7 to the C-terminal receptor binding domain in both systems (Fig. 1A and Table 1). Expression of N-h$\alpha_2$M yielded ~1.6/~3.2 mg per liter of insect and mammalian cell culture, respectively, while the values for C-h$\alpha_2$M were ~1.3/~4.6 mg. N-h$\alpha_2$M formed a dimer of ~170 kDa due to the presence of an intermolecular disulfide bond (C$^{278}$–C$^{431}$), which is also required for dimerization of the authentic full-length protein (Fig. 1A). Consistently, the protein migrated as a monomer of ~85 kDa in the presence of reducing agents. In turn, C-h$\alpha_2$M was monomeric (~75 kDa) and treatment with methylamine did not affect the content of free cysteines or electrophoretic mobility in native-PAGE (Fig. 2B). To follow this up, we qualitatively assayed the content of free sulfhydryl groups by a fluorometric thiol assay kit, which gave a strong fluorescent signal for both the untreated and methylamine-treated C-h$\alpha_2$M samples. This contrasted with native full-length authentic h$\alpha_2$M, which gave no significant signal, and was similar to methylamine-induced authentic h$\alpha_2$M, which likewise gave a strong signal. These assays indicated that the thioester bond was opened in C-h$\alpha_2$M as mentioned above for the full-length recombinant variant, possibly owing to a nucleophilic component of the undisclosed cell-growth medium.

The insect and mammalian systems were also assayed for expression of mature human His$_6$-tagged TGF-$\beta_2$ (Table 1), but without noticeable yields. Therefore, full-length pro-TGF-$\beta_2$ encompassing LAP and mature TGF-$\beta_2$ (see Fig. 1C) was expressed and purified in Expi293F cells as described elsewhere[8], with a final yield of ~2.7 mg and ~2.3 mg of N-terminally octahistidine-tagged and Strep-tagged forms, respectively, per liter of mammalian cell culture (Fig. 2A). The protein migrated as a dimer of ~110 kDa in SEC, which indicates that the characteristic disulfide bonds were formed between the LAP and the mature TGF-$\beta_2$ moieties. The purified protein was partially cleaved before residue A$^{303}$ by host peptidases. Subsequent treatment with the physiological activating endopeptidase furin produced a homogenously cleaved species consisting of LAP associated with the mature cytokine (Fig. 2A). Under physiological conditions, TGF-$\beta_2$ maturation is a complex process that involves a cascade of events under participation of several proteins that interact with the initial complex of pro-TGF-$\beta_2$ and the latent TGF-$\beta$ binding protein (LTBP). LTBP participates as a localizer of pro-TGF-$\beta_2$ to the extracellular matrix, whereas LAP senses the changes and releases mature TGF-$\beta_2$[11]. Previous studies with a Chinese hamster ovary cell expression system benefited from the sensitivity of the LAP domain towards denaturing conditions at very low pH to separate it from mature TGF-$\beta_2$[11,35]. In our case, this was unsuccessful, probably due to different post-translation modifications introduced by Expi293F cells in the highly glycosylated LAP[8,36].

Finally, full-length GRAB was expressed without the cell-wall anchoring region (Fig. 1B) in a bacterial system yielding ~4 mg of pure protein per liter of expression medium after affinity chromatography, IEC and SEC steps (Fig. 2A). The protein migrated as a ~55-kDa species in SEC and as a ~33-kDa species in SDS-PAGE, but the values determined by SEC-MALLS (15.5 kDa; Table 2) were closer to the theoretical mass (15.8 kDa). We attribute this abnormal migration, which was described previously[11], to the highly unstructured character of the protein.

**Interaction analysis of h$\alpha_2$M and GRAB.** Interaction of streptococci with h$\alpha_2$M has been reported to be highly specific[15,16]. Group A, G and C streptococci all bind the native form, whereas only the latter interact with the induced form. This result was attributed to the types of surface proteins, which are specific for each strain. GRAB is found in group A streptococci, and we studied its interaction with native authentic h$\alpha_2$M by surface plasmon resonance. GRAB was immobilized as a ligand through a Strep-tag on a chip with covalently bound streptavidin. In a multicycle experiment, saturation of the ligand was reached with the two highest analyte concentrations, which gave on- and off-rate kinetic constants and results from affinity analysis (Fig. 3A,B). From the sensorgrams during the sequential injections of different analyte concentrations, we observed fast association and slow dissociation of h$\alpha_2$M from GRAB, which indicated stable complex formation. Therefore, the ligand was removed in a regeneration step to make sure that all bound h$\alpha_2$M was eliminated between injections with different analyte concentrations. The group of curves in Fig. 3A,B were fitted to a 1:1 Langmuir interaction model. These calculations revealed a $\chi^2$ value < 10% of R$_{max}$, which is indicative of a good fit. Consistently with the sensorgrams, the association rate constant ($k_a$) and the dissociation rate constant ($k_d$) were $1.32 \times 10^5$ M$^{-1}$s$^{-1}$ and $1.90 \times 10^{-3}$ s$^{-1}$, respectively, with an estimated dissociation halftime ($t_{1/2} = \ln2/k_{off}$) of 365 sec. The equilibrium dissociation constants ($K_D$) from the kinetic and affinity analysis were $1.43 \times 10^{-8}$ M and $3.45 \times 10^{-8}$ M (Tables 3 and 4), respectively, which indicates high affinity and stable complex formation. The complex was also detected by SDS-PAGE and native-PAGE employing fluorophore-labelled GRAB (Fig. 4). Finally, SEC-MALLS analysis (Fig. 3D and Table 2) showed a molecular mass difference of 27.3 kDa over free h$\alpha_2$M, which corresponds to 1.7 molecules of GRAB. Hence, we assume that two molecules of GRAB bind one h$\alpha_2$M tetramer.

Under a similar experimental setup, methylamine-induced authentic h$\alpha_2$M was injected over immobilized GRAB to reach equilibrium and saturation, which enabled analysis by affinity. The affinity data permitted calculation with confidence ($\chi^2 < 10\%$ of R$_{max}$) of the equilibrium dissociation constant ($9.46 \times 10^{-8}$ M), which was three times higher than that of native h$\alpha_2$M (Fig. 3A,B and Table 4). This is consistent with published results, which indicated that GRAB shows preference for native over protease-induced h$\alpha_2$M[16]. The complex was likewise analyzed by SDS-PAGE and native-PAGE with fluorophore-labelled GRAB (Fig. 4). The results showed an increase in the molecular mass of 26.5 kDa over noncomplexed induced h$\alpha_2$M, which is equivalent to the results for native h$\alpha_2$M.

To map down the region of h$\alpha_2$M engaged in GRAB binding, we repeated the above experiments with N-h$\alpha_2$M and C-h$\alpha_2$M. In a similar multicycle experimental setup, we could not detect any interaction. However, previous incubation of the proteins at 37 °C for one hour apparently enabled complex formation. Protein remained complexed over time after injection and washing of the chip (Fig. 3B), but in this case we could not determine the affinity constants due to the experimental setup. The complexes were subsequently evaluated in native-PAGE using fluorophore-labelled GRAB (Fig. 3D). In this case, we detected the interaction of GRAB with C-h$\alpha_2$M but not with N-h$\alpha_2$M.

**Interaction analysis of h$\alpha_2$M and pro-TGF-$\beta_2$.** Previous biochemical data had revealed that h$\alpha_2$M binds TGF-$\beta_2$ mainly through a mature cytokine segment spanning residues A$^{343}$-Y$^{367}$, in which W$^{354}$ plays a major role[10]. No data have been reported on the role of LAP. However, inspection of the crystal structure of homologous pro-TGF-$\beta_1$ (see Protein Data Bank code 3RJR[37]) reveals that the interacting segment is partially shielded by LAP. Other studies employing a library of overlapping glutathione *S*-transferase fusion proteins ascribed the potential binding site for TGF-$\beta_1$ to segment V$^{723}$-T$^{761}$ of h$\alpha_2$M[38], which was subsequently narrowed down to E$^{737}$-V$^{756}$ employing synthetic peptides[39]. However, further details on the mechanism are unknown. To further shed light, we set out to characterize binding of pro-TGF-$\beta_2$ to h$\alpha_2$M. We checked the interaction by surface plasmon resonance in multicycle experiments with immobilized pro-TGF-$\beta_2$ as ligand but could not detect complex formation. Only after analysis by native-PAGE using fluorophore-labelled pro-TGF-$\beta_2$ we observed interaction with native authentic h$\alpha_2$M but not with the induced form or the short variants (Fig. 4). Given that the pro-TGF-$\beta_2$ sample contained a mixture of cleaved and intact protein, we assayed N-terminally His$_6$-tagged pro-TGF-$\beta_2$ with native h$\alpha_2$M in native-PAGE followed by Western blotting. The two proteins were not co-migrating (data not shown). Thus, we conclude that LAP prevents h$\alpha_2$M from binding mature TGF-$\beta_2$ as suggested by structural studies on pro-TGF-$\beta_1$.

**Conclusions.** Protein h$\alpha_2$M is a sophisticated player to spatially and temporally restrict and regulate key physiological processes that control the distribution and activity of many proteins, including peptidases, cytokines, hormones and other physiological effectors[1]. Since the 1940s, several efforts have been made to understand its mechanism of action *in vivo* and *in vitro*, but they have been hampered by the unavailability of high-yield recombinant expression systems. Here, we developed insect and mammalian systems for the full-length protein and shorter fragments. The former was mainly produced in an induced state, possibly due to media components that cause induction during the time scale of expression. Thus, other media with a regulated composition will be assayed to reevaluate the recombinant systems.

The recombinant proteins *plus* authentic h$\alpha_2$M were analyzed for binding with GRAB and pro-TGF-$\beta_2$. The former tightly bound native and methylamine-induced authentic h$\alpha_2$M, with two molecules of GRAB per h$\alpha_2$M tetramer. The short variants, especially C-h$\alpha_2$M, likewise complexed GRAB, but apparently through a different mechanism from the full-length forms. In contrast, full-length pro-TGF-$\beta_2$ did not complex any h$\alpha_2$M variant, probably owing to steric hindrance by the N-terminal LAP domain.

## References

1. Garcia-Ferrer, I., Marrero, A., Gomis-Rüth, F. X. & Goulas, T. $\alpha_2$-Macroglobulins: structure and function. *Subcell. Biochem.* **83**, 149–183 (2017).
2. Goulas, T. *et al.* Structural and functional insight into pan-endopeptidase inhibition by $\alpha_2$-macroglobulins. *Biol. Chem.* **398**, 975–994 (2017).
3. Barrett, A. J. & Starkey, P. M. The interaction of $\alpha_2$-macrolgobulin with proteinases. *Biochem. J.* **133**, 709–724 (1973).
4. Sottrup-Jensen, L., Sand, O., Kristensen, T. & Fey, G. H. The $\alpha$-macroglobulin bait region. Sequence diversity and localisation of cleavage sites for proteinases in five mammalian $\alpha$-macroglobulins. *J. Biol. Chem.* **264**, 15781–15789 (1989).
5. Garcia-Ferrer, I. *et al.* Structural and functional insights into *Escherichia coli* $\alpha_2$-macroglobulin endopeptidase snap-trap inhibition. *Proc. Natl. Acad. Sci. USA* **112**, 8290–8295 (2015).
6. Marrero, A. *et al.* The crystal structure of human $\alpha_2$-macroglobulin reveals a unique molecular cage. *Angew. Chem. Intl. Ed.* **51**, 3340–3344 (2012).
7. Bieth, J. G., Tourbez-Perrin, M. & Pochon, F. Inhibition of $\alpha_2$-macroglobulin-bound trypsin by soybean trypsin inhibitor. *J. Biol. Chem.* **256**, 7954–7957 (1981).
8. del Amo-Maestro, L., Marino-Puertas, L., Goulas, T. & Gomis-Rüth, F. X. Recombinant production, purification, crystallization, and structure analysis of human transforming growth factor $\beta$2 in a new conformation. *Sci. Rep.* **9**, 8660 (2019).
9. LaMarre, J., *et al.* An $\alpha_2$-macrolgobulin receptor-dependent mechanism for the plasma clearance of transforming growth factor-$\beta$1 in mice. *J. Clin. Invest.* **87**, 39–44 (1991).
10. Liu, Q. *et al.* Identification of the high affinity binding site in transforming growth factor-$\beta$ involved in complex formation with $\alpha_2$-macroglobulin. Implications regarding the molecular mechanisms of complex formation between $\alpha_2$-macroglobulin and growth factors, cytokines, and hormones. *J. Biol. Chem.* **276**, 46212–46218 (2001).
11. Annes, J. P., Munger, J. S. & Rifkin, D. B. Making sense of latent TGF-$\beta$ activation. *J. Cell Sci.* **116**, 217–224 (2003).
12. O'Connor-McCourt, M. D. & Wakefield, L. M. Latent transforming growth factor-$\beta$ in serum. A specific complex with $\alpha_2$-macroglobulin. *J. Biol. Chem.* **262**, 14090–14099 (1987).
13. Webb, D. J., Wen, J., Karns, L. R., Kurilla, M. G. & Gonias, S. L. Localisation of the binding site for the transforming growth factor-beta in human $\alpha$2-macroglobulin to a 20-kDa peptide that also contains the bait region. *J. Biol. Chem.* **273**, 13339–13346 (1998).
14. Arandjelovic, S., van Sant, C. L. & Gonias, S. L. Limited mutations in full-length tetrameric human $\alpha_2$-macroglobulin abrogate binding of platelet-derived growth factor-BB and transforming growth factor-$\beta$1. *J. Biol. Chem.* **281**, 17061–17068 (2006).
15. Godehardt, A., Hammerschmidt, S., Frank, R. & Chhatwal, G. S. Binding of $\alpha_2$-macroglobulin to GRAB (Protein G-related $\alpha_2$-macroglobulin binding protein), an important virulence factor of group A streptococci, is mediated by two charged motifs in the $\Delta$A region. *Biochem. J.* **381**, 877–885 (2004).
16. Rasmussen, M., Müller, H.-P. & Björck, L. Protein GRAB of *Streptococcus pyogenes* regulates proteolysis at the bacterial surface by bnding $\alpha_2$-macroglobulin. *J. Biol. Chem.* **274**, 15336–15344 (1999).
17. Goulas, T., Garcia-Ferrer, I., Garcia-Pique, S., Sottrup-Jensen, L. & Gomis-Rüth, F. X. Crystallization and preliminary X-ray diffraction analysis of eukaryotic $\alpha_2$-macroglobulin family members modified by methylamine, proteases and glycosidases. *Molec. Oral Microbiol.* **29**, 354–364 (2014).
18. Andersen, G. R. *et al.* Crystallisation of proteins of the $\alpha_2$-macroglobulin superfamily. *Ann. N.Y. Acad. Sci.* **737**, 444–446 (1994).
19. Goulas, T. *et al.* The pCri System: a vector collection for recombinant protein expression and purification. *PloS one* **9**, e112643 (2014).
20. Hemsley, A., Arnheim, N., Toney, M. D., Cortopassi, G. & Galas, D. J. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucl. Acids Res.* **17**, 6545–6551 (1989).
21. Xie, Q. *et al.* TubeSpin bioreactor 50 for the high-density cultivation of Sf-9 insect cells in suspension. *Biotechnol. Lett.* **33**, 897–902 (2011).
22. Strober, W. Trypan blue exclusion test of cell viability. *Curr. Protoc. Immunol.* **21**, A.3B.1–A.3B.2 (2001).
23. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166**, 557–580 (1983).
24. Sottrup-Jensen, L., Petersen, T. E. & Magnusson, S. A thiol-ester in $\alpha_2$-macroglobulin cleaved during proteinase complex formation. *FEBS Lett.* **121**, 275–279 (1980).
25. Imber, M. J. & Pizzo, S. V. Clearance and binding of two electrophoretic "fast" forms of human $\alpha_2$-macroglobulin. *J. Biol. Chem.* **256**, 8134–8139 (1981).
26. Morrot, A. *et al.* Human T cell responses against the major cysteine proteinase (cruzipain) of *Trypanosoma cruzi*: role of the mutlifunctional $\alpha_2$-macroglobulin receptor in antigen presentation by monocytes. *Int. Immunol.* **9**, 825–834 (1997).
27. Haider, S. R., Sharp, B. L. & Reid, H. J. A comparison of Tris-glycine and Tris-tricine buffers for the electrophoretic separation of major serum proteins. *J. Separat. Sci.* **34**, 2463–2467 (2011).
28. Sinha-Datta, U., Khan, S. & Wadgaonkar, D. Label-free interaction analysis as a tool to demonstrate biosimilarity of therapeutic monoclonal antibodies. *Biosimilars* **15**, 83–91 (2015).
29. Zhao, J. *et al.* Novel method for measurement of heparin anticoagulant activity using SPR. *Anal. Biochem.* **526**, 39–42 (2017).
30. Shen, X., Hacker, D. L., Baldi, L. & Wurm, F. M. Virus-free transient protein production in Sf9 cells. *J. Biotechnol.* **171**, 61–70 (2014).
31. Sottrup-Jensen, L. $\alpha$-Macroglobulin: structure, shape, and mechanism of proteinase complex formation. *J. Biol. Chem.* **264**, 11539–11542 (1989).
32. van Rompaey, L. & Marynen, P. Temperature-dependent biosynthesis of thiol esters in baculovirus recombinant $\alpha_2$M and PZP. *Ann. N. Y. Acad. Sci.* **737**, 506–509 (1994).
33. Baxter, R. H. G. *et al.* Structural basis for conserved complement factor-like function in the antimalarial protein TEP1. *Proc. Natl. Acad. Sci. USA* **104**, 11615–11620 (2007).

34. Grøn, H., Thøgersen, I. B., Enghild, J. J. & Pizzo, S. V. Structural and functional analysis of the spontaneous reformation of the thiol ester bond in human $\alpha_2$-macroglobulin, rat $\alpha_1$-inhibitor-3 and chemically modified derivatives. *Biochem. J.* **318**, 539–545 (1996).
35. Zou, Z. & Sun, P. D. An improved recombinant mammalian cell expression system for human transforming growth factor-$\beta$. *Prot. Expr. Purif.* **50**, 9–17 (2006).
36. Croset, A. *et al.* Differences in the glycosylation of recombinant proteins expressed in HEK and CHO cells. *J. Biotechnol.* **61**, 336–348 (2012).
37. Shi, M. *et al.* Latent TGF-$\beta$ structure and activation. *Nature* **474**, 343–349 (2011).
38. Webb, D. J., Roadcap, D. W., Dhakephalkar, A. & Gonias, S. L. A 16-amino acid peptide from human $\alpha$2-macroglobulin binds transforming growth factor-$\beta$ and platelet-derived growth factor-BB. *Prot. Sci.* **9**, 1986–1992 (2000).
39. Arandjelovic, S., Freed, T. A. & Gonias, S. L. Growth factor-binding sequence in human $\alpha_2$-macroglobulin targets the receptor-binding site in transforming growth factor-$\beta$. *Biochemistry* **42**, 6121–6127 (2003).

## Acknowledgements

## Author Contributions

T.G. and F.X.G.R. conceived and supervised the work; L.M.P. and L.d.A.M. produced and purified the protein; L.M.P. and L.d.A.M. performed biochemical studies; M.T.M. guided the Biacore experiments; and T.G. and F.X.G.R. wrote the paper with contributions from all authors.

## Additional Information

**Supplementary information** accompanies this paper at https://doi.org/10.1038/s41598-019-45712-z.

**Competing Interests:** The authors declare no competing interests.

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