

# The IgGFc-binding protein FCGBP is secreted with all GDPH sequences cleaved but maintained by interfragment disulfide bonds

Received for publication, February 25, 2021, and in revised form, May 30, 2021. Published, Papers in Press, June 11, 2021,

<https://doi.org/10.1016/j.jbc.2021.100871>

Erik Ehrencrona, Sjoerd van der Post<sup>1</sup>, Pablo Gallego, Christian V. Recktenwald<sup>1</sup>, Ana M. Rodriguez-Pineiro, Maria-Jose Garcia-Bonete, Sergio Trillo-Muyo, Malin Bäckström<sup>1</sup>, Gunnar C. Hansson<sup>1</sup>, and Malin E. V. Johansson<sup>\*1</sup>

From the Department of Medical Biochemistry, Institute of Biomedicine, University of Gothenburg, Gothenburg, Sweden

Edited by Gerald Hart

Mucus forms an important protective barrier that minimizes bacterial contact with the colonic epithelium. Intestinal mucus is organized in a complex network with several specific proteins, including the mucin-2 (MUC2) and the abundant IgGFc-binding protein, FCGBP. FCGBP is expressed in all intestinal goblet cells and is secreted into the mucus. It is comprised of repeated von Willebrand D (vWD) domain assemblies, most of which have a GDPH amino acid sequence that can be autocatalytically cleaved, as previously observed in the mucins MUC2 and mucin-5AC. However, the functions of FCGBP in the mucus are not understood. We show that all vWD domains of FCGBP with a GDPH sequence are cleaved and that these cleavages occur early during biosynthesis in the endoplasmic reticulum. All cleaved fragments, however, remain connected *via* a disulfide bond within each vWD domain. This cleavage generates a C-terminal-reactive Asp-anhydride that could react with other molecules, such as MUC2, but this was not observed. Quantitative analyses by MS showed that FCGBP was mainly soluble in chaotropic solutions, whereas MUC2 was insoluble, and most of the secreted FCGBP was not covalently bound to MUC2. Although FCGBP has been suggested to bind immunoglobulin G, we were unable to reproduce this binding *in vitro* using purified proteins. In conclusion, while the function of FCGBP is still unknown, our results suggest that it does not contribute to covalent crosslinking in the mucus, nor incorporate immunoglobulin G into mucus, instead the single disulfide bond linking each fragment could mediate controlled dissociation.

Intestinal mucus is part of our innate defense against microbes and other hazards such as chemical challenges and mechanical stress from dehydration or ingested particles. The secreted mucus forms an organized structure with mucin-2 (MUC2) as the main structural component. Mucus also contains a number of other molecules, some from shed cells and others secreted by epithelial cells (1, 2). The core mucus proteins are produced and secreted by goblet cells, of which

the calcium-activated chloride channel 1 and the IgGFc-binding protein (FCGBP) are as abundant in the intestinal mucus as MUC2 (3). The secreted mucus expands and forms a net-like structure building the inner mucus layer that keeps bacteria at a distance, but molecular details are largely missing (1). Calcium-activated chloride channel 1 is, in contrast to its name, a metalloprotease controlling further mucus expansion in the colon in a similar manner as meprin  $\beta$  does in the small intestine (4, 5). However, the role of the FCGBP molecule is still a conundrum.

Human FCGBP is made up of 13 serially repeated von Willebrand D (vWD) domains, whereas mouse *Fcgbp* contains 7. This is more vWD domains than in any other known vertebrate protein (6), and human FCGBP has more than three times as many vWDs as the von Willebrand factor (vWF) protein, for which these domains were named. The secreted mucins MUC2, mucin-5AC (MUC5AC), mucin-5B, and mucin-6 all have repeated vWD domains just like the vWF, tectorins, subcommissural organ-spondins, otogelins, and zonadhesin, although these vWD sequences cluster separately in a phylogenetic tree (6). The central proline-, serine-, and threonine-rich domains of mucins are highly glycosylated, in contrast to FCGBP that only contains vWD domains. The MUC2 and MUC5AC mucins have one vWD domain in their C-terminus containing a Gly-Asp-Pro-His (GDPH) amino acid sequence cleaved by an autocatalytic process (7, 8). The GDPH sequence is also found in vWD domains of other proteins (9–12) and in proteins without vWD domains (13–15). Interestingly, the human FCGBP has this motif embedded within 11 of its 13 vWDs and the mouse *Fcgbp* within five of its seven vWDs (16, 17). The Asp-Pro bond is known to be unstable *in vitro* based on previous observations. These motifs are likely autocatalytically cleaved, hydrolyzing the peptide bond between the amino acids Asp and Pro (11, 15, 18), generating a reactive and unstable Asp-anhydride in the newly generated C terminus (13, 19, 20). No real functional relevance of this cleavage has been observed for most proteins, but cleavage could be important for correct folding and potentially calcium dependent (9, 12, 15). Binding to the Asp-anhydride in the new C terminus, generated through GDPH autocatalytic

\* For correspondence: Malin E.V. Johansson, [malin.johansson@medkem.gu.se](mailto:malin.johansson@medkem.gu.se)

## GDPH cleavage in FCGBP

cleavage, has been observed for a few proteins such as the inter-alpha-trypsin inhibitor heavy chain 3 (ITH3) that binds to chondroitin sulfate on bikunin to form the pre-alpha-inhibitor (21, 22). The Asp on ITH3 is linked to the C6 position of a GalNAc of the chondroitin sulfate as also observed for related molecules (19, 20). The many GalNAc residues on MUC2 suggest that the GDPH motifs potentially mediate the previously observed covalent coupling of FCGBP to MUC2 (17). Formation of an isopeptide bond to Lys has also been proposed (15) similar to isopeptide cross-linking by transamidation responsible for stabilizing the mucus (23).

The cleavage of ITH3 is promoted by the low pH in the late secretory pathway as observed for MUC2 (8, 24). In contrast, the GDPH autocatalytic cleavage in the vWD-containing mucin MUC5AC occurs during early biosynthesis (7). FCGBP is named based on the original observation that it binds immunoglobulin G (IgG), something that has been suggested to be utilized as treatment to trap human immunodeficiency virus in cervical mucus (16, 25–28). It has also been implicated to bind another mucus component, trefoil factor 3 (29). A functional role in cross-linking mucus or incorporating components into the mucus is thus proposed for FCGBP, but further studies are required to identify how FCGBP can contribute to mucosal protection.

Here, we sought to address the extent of autocatalytic cleavage and structural significance of the GDPH motif in mouse and human FCGBP. In addition to characterizing the protein processing, we further investigated the previously described interactions of FCGBP with MUC2 and IgG.

## Results

### Human and mouse FCGBP domain structure and expression

Human FCGBP (5405 amino acids) contains 13 vWD domains, while *Fcgbp* in mouse (2583 amino acids), as well as in other rodents, is comprised of seven vWD domains (Fig. 1A). The N-terminal, non-vWD-containing domain is longer in humans, as well as in many other species, than in mouse and other rodents (6). To investigate the features of the different vWD domain sequences in mouse and human FCGBP, the sequences were aligned (Fig. S1) and a phylogenetic tree analysis was performed. The sequences of all vWD assemblies were used, highlighting homogeneity between domains within the molecules and cross-species (Fig. 1B). This showed that the human FCGBP has three consecutive vWD triplets in the central region (vWDs, 3–11) (16), whereas the mouse only has one of these (vWDs, 3–5). The vWD3, vWD6, and vWD9 differ the most, and these domains have additional shorter sequences interspersed in the 8 cysteine domain and extended sequences after both trypsin inhibitor-like domain and E domains (Fig. S1). The vWD1 and vWD2 cluster together as do the vWD12 in human and vWD6 in mouse. The last vWD (13 in humans and seven in mice) cluster together but separate from the others, suggesting larger sequence difference and possibly a specific function. The vWD13 and the N-terminal domain of the human sequence were selected as unique domains suitable for generating FCGBP antibodies.

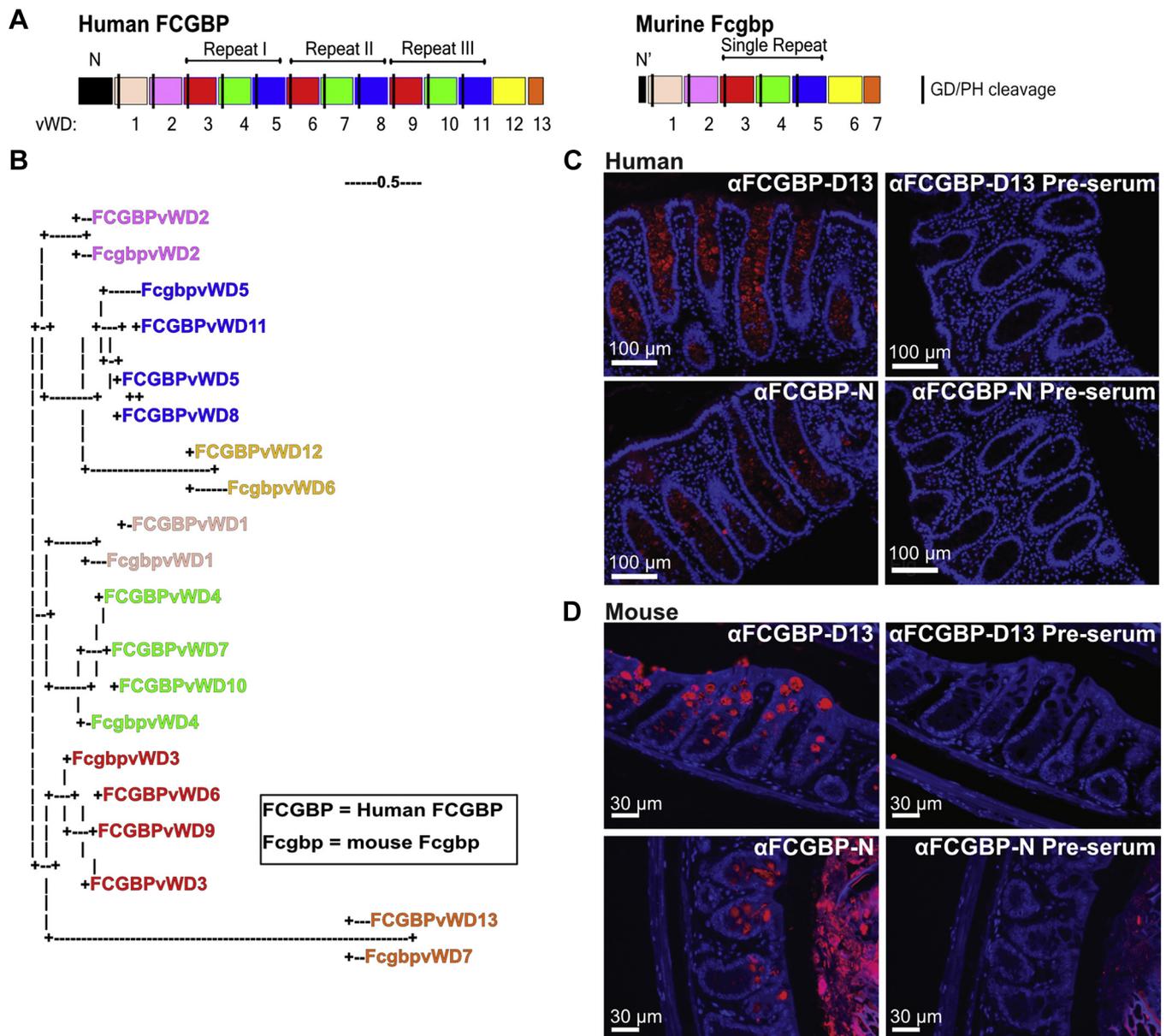
Recombinantly expressed FCGBP-N and FCGBP-D13 were used to generate the rabbit polyclonal antisera,  $\alpha$ FCGBP-N and  $\alpha$ FCGBP-D13, respectively. Immunostaining with both antisera detected FCGBP throughout the full length of the crypt in Carnoy-fixed sections from human sigmoid colon using the preimmunization serum as the control (Fig. 1C). The same was true for stained sections from mouse distal colon. Mouse *Fcgbp* was also detected along the full length of the crypt, with a strong difference in intensity between pre-immunized and immunized serum for both the  $\alpha$ FCGBP-D13 and  $\alpha$ FCGBP-N (Fig. 1D). This indicates that the full protein is expressed in intestinal goblet cells and mainly detected in the granules of the goblet cell theca.

### FCGBP is autocatalytically cleaved but connected via disulfide bridges

FCGBP abundance was investigated in mucus and cells from human and mouse tissues by SDS-PAGE and Western blot (WB) with antibodies detecting different parts of FCGBP (Fig. 2E). In nonreduced samples, both the  $\alpha$ FCGBP-N and  $\alpha$ FCGBP-D13 antisera detected a >460-kDa band in humans (Fig. 2A) and a >268-kDa band in mice (Fig. 2B). The size of these bands corresponded to the full-length FCGBP in humans and mice, respectively. The specificity of the antisera was verified as *Fcgbp* was not detected in lysates from *Fcgbp*<sup>-/-</sup> mice (KO). The content of bands of equivalent sizes was analyzed by MS, identifying peptides through all the domains of the proteins in both species except for the small mouse N'-GD fragment in the lysate (Fig. 2F).

Upon reduction, smaller fragments were detected, showing that FCGBP was cleaved yet held together by disulfide bonds in both humans and mice (Fig. 2, C and D). A band migrating around 100 kDa was detected using  $\alpha$ FCGBP-D13 in both human and mouse cell lysate that was absent in *Fcgbp*<sup>-/-</sup> mice. In both species, this band was also identified by the  $\alpha$ FCGBP-3517 antibody, which is in accordance with the putative molecular mass (98 kDa) of the C-terminal fragment after GDPH cleavage in the vWD11 (human) or vWD5 (mouse) (Fig. 2E). Indeed, MS analysis of these bands verified that most FCGBP peptides originate from this fragment (Fig. 2F).

The  $\alpha$ MUC2C3 antiserum was raised to a peptide on the C-terminal side of the cleaved GDPH in human MUC2. This antiserum showed cross-reactivity with both reduced and nonreduced FCGBP and likely detects most FCGBP vWD domains in both humans and mice (Fig. 2, B–E). The specific bands for mouse *Fcgbp* were revealed using tissue from both WT and *Muc2*<sup>-/-</sup>, and they were not present in *Fcgbp*<sup>-/-</sup> mice (Fig. 2D). As expected, the  $\alpha$ MUC2C3 detected bands in the 40- to 55-kDa range in both humans and mice, corresponding to the fragments generated after cleavage at GDPH of two adjacent vWD domains. The antibodies  $\alpha$ FCGBP-3517 and  $\alpha$ FCGBP-3564 also detect bands of this size range, at least in human samples.  $\alpha$ FCGBP-3564 likely detects the longer human N-terminal domain, while  $\alpha$ FCGBP-3517 can detect the cleaved vWD5 and vWD8 domains. The identity of these bands, as revealed by MS analysis, was less conclusive because



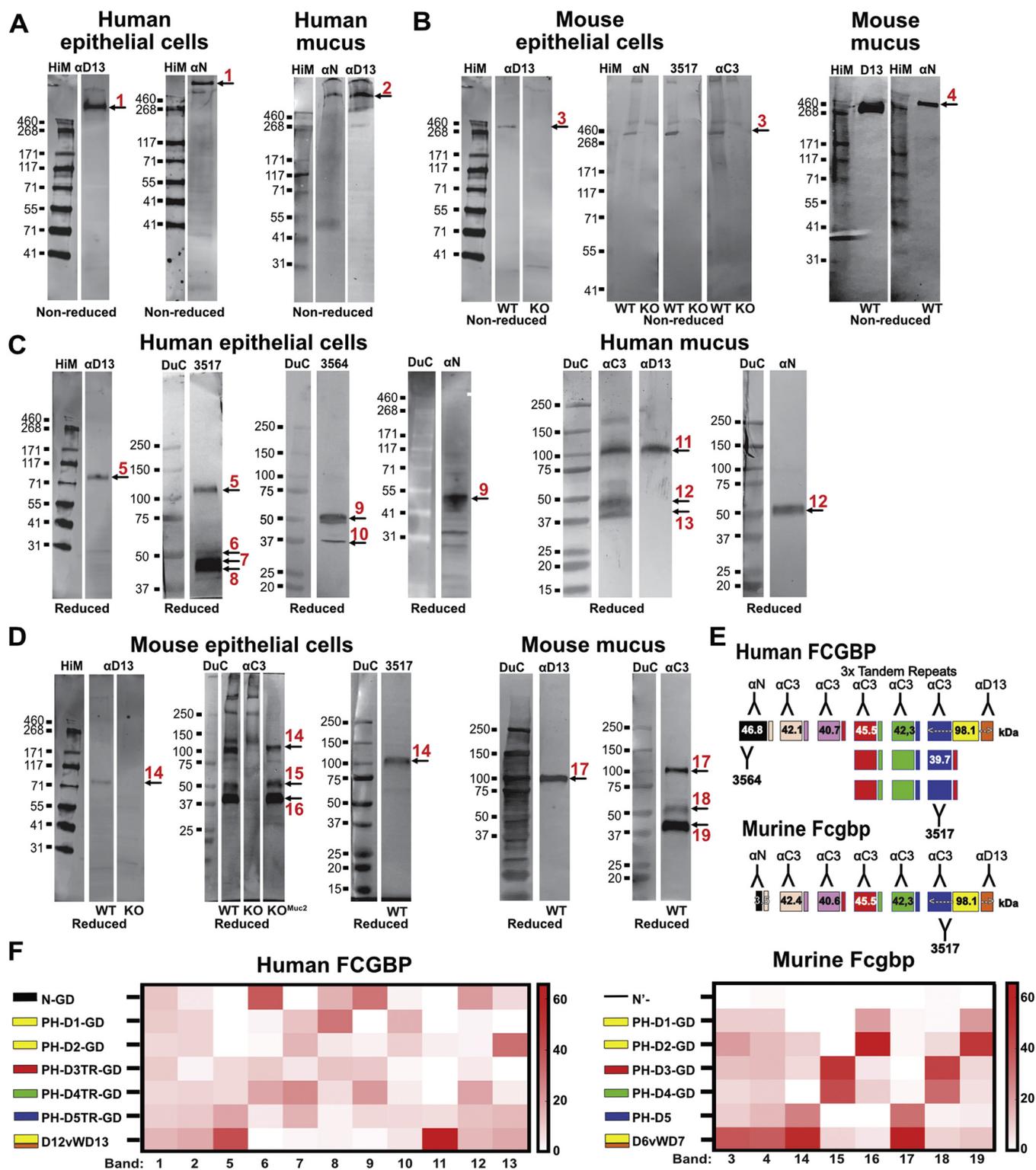
**Figure 1. The vWD domain-rich FCGBP is expressed by all intestinal goblet cells.** *A*, illustration of the domain composition of human and mouse FCGBP. The human N-terminal domain is extended compared with the mouse variants. Each D-assembly is in itself constituted by a vWD followed by a C8 and TIL domain. All GDPH sites are indicated, and domains are color-coded. *B*, alignments of the vWD domains from mouse and human show that the D-modules 3 to 11 are three repeated units in humans, while the D3 to D5 constitute a single repeat in mouse Fcgbp. A phylogenetic tree with the vWD domains of human FCGBP and mouse Fcgbp. The domains are color-coded as in panel *A* and show similar organization in mice and humans with a repeated central part in human FCGBP. *C* and *D*, immunostained Carnoy-fixed 4- $\mu$ m sections from human sigmoid (*C*) and mouse distal colon (*D*). Sections are stained with both  $\alpha$ FCGBP-N and  $\alpha$ FCGBP-D13 antiserum against FCGBP (red). Sections stained with preimmunization serum ( $\alpha$ FCGBP-N,  $\alpha$ FCGBP-D13 pre-serum) are presented as controls. Hoechst-stained nuclei are shown in blue. C8, 8 cysteine; FCGBP, IgGfC-binding protein; GDPH, Gly-Asp-Pro-His; TIL, trypsin inhibitor-like domain; vWD, von Willebrand D.

of the repetitive nature of the human sequence and that all fragments are present in the gel enhancing the risk of contamination between bands. Identification of these bands in the mouse samples was more definite because vWD1 and vWD2 correspond to smaller sizes, whereas vWD3 and vWD4 are present in the bands corresponding to larger size molecules. The GDPH-cleaved N-terminal part of mouse Fcgbp was mostly not detected by MS, likely because of its predicted small size (6 kDa) with few potential tryptic peptides. In human samples, the N-terminal part was detected as a band at

almost 55 kDa (predicted to be 46.8 kDa) as observed with both  $\alpha$ FCGBP-N and  $\alpha$ FCGBP-3564, suggesting that it has numerous post-translational modifications. A smaller 37-kDa band was sometimes observed in human samples identified as fragments of vWD1 and vWD2 (Fig. 2, *C* and *E*).

Analysis of data from the mucus proteome from mice and humans (2, 3) revealed semitryptic peptides derived from the PH side of the GDPH motif after cleavage. For human FCGBP, peptides from all sites except for *PHYTTFDGHR* in vWD9 were identified (Table 1). In mouse mucus, all PH peptides

## GDPH cleavage in FCGBP



**Figure 2. FCGBP is cleaved at the GDPH sequences and associated as a full molecule by disulfide bonds.** *A*, epithelial cell lysate and mucus from human sigmoid biopsies analyzed by SDS-PAGE under nonreducing conditions with WB using both αN and αD13 FCGBP antisera. A band >460 kDa was detected. *B*, analysis as in panel *A* of samples from the distal colon of WT and *Fcgbp*<sup>-/-</sup> (KO) mice. Additional antiserum, αFCGBP-3517 (3517), and the cross-reacting αMUC2C3 (αC3) also specifically detected the intact Fcgbp with a size >268 kDa. The leftmost marker in panels *A* and *B* are identical as human and mouse samples were run on the same gel. This is also the case in panels *C* and *D*. *C*, reduced human sigmoid epithelial cell lysate and mucus analyzed by SDS-PAGE and WB using different antibodies against FCGBP as indicated detected smaller fragments. The sizes fit well with estimated sizes around 50 kDa for all fragments except the longer C terminus with a size >100 kDa. *D*, analysis as in panel *C* of reduced mouse samples revealing fragments as predicted around 50 kDa and a larger C-terminal fragment (>100 kDa). *E*, schematic presentation of the preferred binding of the different antibodies. *F*, Coomassie Brilliant Blue–stained gel bands of equivalent sizes are marked with arrows, and numbers in panels *A–D* were analyzed through in-gel digestion and MS. The percentage of total FCGBP/Fcgbp peptides originating from each of the putative GDPH cleavage products is shown in a heatmap. FCGBP, IgGf-binding protein; GDPH, Gly-Asp-Pro-His; WB, Western blot.

**Table 1**  
Identified FCGBP GD/PH peptides in human colonic mucus<sup>a</sup>

vWD	Start	End	Peptide sequence	Observed <i>m/z</i>	Charge [H] <sup>+</sup>	Molecular weight calculated	$\Delta$ ppm
1	478	486	PHYTFDGR	547.2585	2	1092.4989	3
3	1257	1266	PHYTFDGR				
6	2457	2466	PHYTFDGR				
2	870	878	PHYVSFDGR	359.8433	3	1076.504	4
4	1679	1687	PHYHSFDGR	372.5055	3	1114.4944	0
7	2880	2888	PHYHSFDGR				
10	4081	4089	PHYHSFDGR				
5	2078	2087	PHYVTLDGHR	597.8052	3	1193.5942	1
8	3278	3287	PHYVTLDGHR				
11	4480	2289	PHYVTLDGHR				
9	3660	3669	PHYTFDGR	Not observed			

<sup>a</sup> No modifications were observed. False discovery rate was <1% for both peptides and proteins.

were detected except PHYNSFDGWSFDQGTCTNYL-LAGTLCPGVNAEGLTPFTVTTK in vWFD4 (Table 2), likely missed in the MS analysis because of its large size. Together the results suggest that both mouse and human FCGBP in both cells and mucus are cleaved at all GDPH sequences and held together by disulfide bonds.

### The Fcgbp signal sequence cleavage site

To identify the N-terminal end after signal peptidase cleavage, a nonreduced band from purified recombinant mouse Fcgbp or lysate from human samples were analyzed by MS. A tryptic peptide SCQGIQCASGQR corresponding to removal of the first 26 amino acids was observed for mouse and the peptide ASVDLK corresponding to removal of the first 23 amino acids for human FCGBP (Fig. 3A). In mouse, this corresponds to the cleavage site predicted by *in silico* analysis (Fig. S2A). Prediction of the signal peptidase cleavage site in the human sequence was less conclusive with a main predicted site after amino acid 21, but additional sites can be possible, and combined with our MS analysis of the N-terminal peptide, cleavage in position 23 is likely favored *in vivo* (Fig. S2B).

### Recombinant FCGBP is GDPH-cleaved in the endoplasmic reticulum and held together by disulfide bonds

The mouse Fcgbp was recombinantly expressed in Chinese hamster ovary (CHO) cells, and when analyzed by denaturing gel electrophoresis, it migrated with an estimated size >268 kDa (Fig. 3B), similarly to Fcgbp in mucus and epithelial cells. Reduction also dissociated this molecule into smaller fragments. Thus, the recombinantly expressed protein from non-mucus-producing cells was similar to the *in vivo*-produced protein with autocatalytic GDPH cleavages. That the

GDPH cleavages were not introduced by heating the sample was tested, and cleaved protein fragments were present also without boiling the protein in the sample buffer (Fig. 3C).

To address when Fcgbp was autocatalytically cleaved during passage through the secretory pathway, lysates of Fcgbp-expressing cells were treated with endoglycosidase H (EndoH). This enzyme cleaves off *N*-glycans of the high mannose form present only on proteins in the endoplasmic reticulum. EndoH treatment resulted in a faster migrating C-terminal FCGBP fragment (Fig. 3D). This Fcgbp fragment is predicted to contain four potential *N*-glycans (Fig. S2C), and the shift in size predicts that Fcgbp is cleaved in the endoplasmic reticulum, especially observed when accumulated because of overexpression. This suggests that the GDPH cleavage takes place early during biosynthesis, likely in the endoplasmic reticulum.

To ensure the cleavage to occur in the GDPH motif, a shorter recombinant protein containing only the first two human vWD domains (FCGBP-ND2) was expressed and the products were separated by SDS-PAGE. Three bands could be observed (Fig. 3E). The bands were cut out, and the putative new N termini were labeled with formaldehyde followed by in-gel digestion performed with the proteases AspN or trypsin (cleave as ArgC after dimethylation) and analyzed by MS. The upper band (band 1) mostly contained peptides from PH-D2-GD, the middle band (band 2) covered the N-GD part, and the lower band (band 3) contained peptides from PH-D1. This confirmed that the FCGBP-ND2 was cleaved at both GDPH cleavage sites.

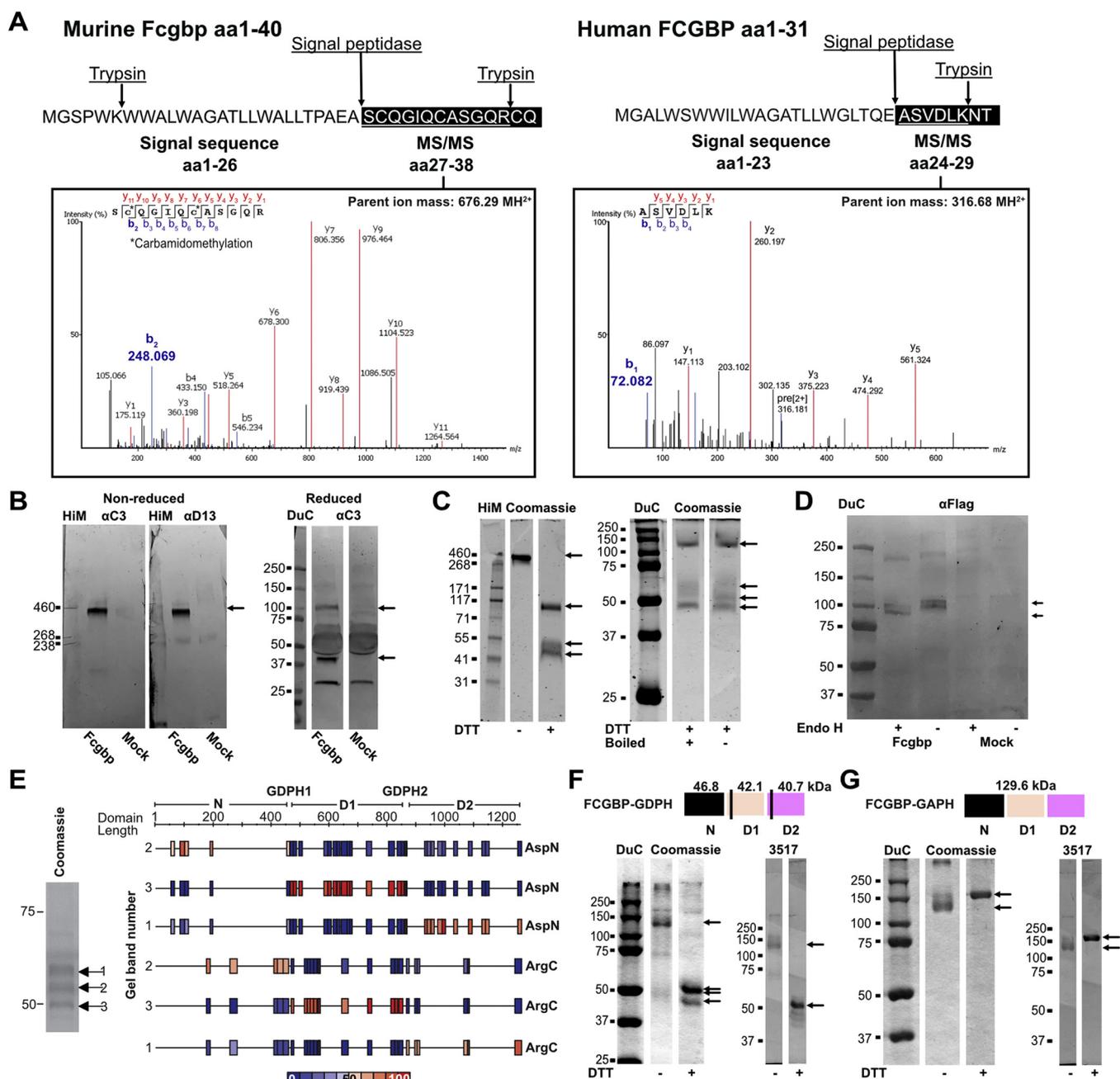
The FCGBP-ND2 protein migrated as one band around 130 kDa under nonreducing conditions and again produced smaller fragments upon reduction (Fig. 3F). When the Asp residue of the two GDPH sequences were mutated to Ala and

**Table 2**  
Identified Fcgbp GD/PH cleaved peptides in mouse colonic mucus<sup>a</sup>

vWD	Start	End	Peptide sequence	Observed <i>m/z</i>	Charge [H] <sup>+</sup>	Molecular weight calculated	$\Delta$ ppm
1	62	70	PHYTFDGR	365.1736	3	1092.4988	0
2	454	462	PHYVSFDGK	350.5067	3	1048.4978	0
3	842	850	PHYVTFDGR	546.2675	2	1090.5195	1
4	1259	1300	PHYNSFDGWSFDQGTCTNYLLAGTL CPGVNAEGLTPFTVTTK	Not observed			
5	1658	1667	PHYVTLDGHR	598.7936	2	1195.5734	1

<sup>a</sup> No modifications were observed. False discovery rate was <1% for both peptides and proteins.

## GDPH cleavage in FCGBP



**Figure 3. FCGBP is cleaved at the signal peptide and GDPH sequences early during biosynthesis.** *A*, a semitryptic peptide identifies the signal peptidase cleavage in nonreduced purified recombinant mouse Fcgbp (left) and human FCGBP (right) with corresponding MS/MS spectra shown. *B*, recombinant Fcgbp produced and secreted by a stable Chinese hamster ovary (CHO)-K1 clone and a mock control analyzed by WB under nonreducing (left) and reducing (right) conditions. The blots are stained by  $\alpha$ MUC2C3 ( $\alpha$ C3) and  $\alpha$ FCGBP-D13 ( $\alpha$ D13), and identified bands correspond in size to the band in mucus samples. *C*, FLAG purification of recombinant Fcgbp on Coomassie Brilliant Blue-stained gels show bands matching WB results in panel *B* (left), and reduction (not boiling) dissociates the fragments (right). *D*, lysates from CHO-K1 cells expressing recombinant Fcgbp were subjected to endoglycosidase H treatment. Treated and nontreated samples along with corresponding controls from mock lysates were analyzed by WB and stained with  $\alpha$ FLAG. A shift in the C-terminal fragments was observed, indicating GDPH cleavage to occur in the endoplasmic reticulum containing high mannose N-glycans. *E*, reduced FCGBP-ND2 separated on SDS-PAGE into three bands that were digested in-gel at Asp (AspN) or Arg (ArgC) and analyzed by MS. The location in the recombinant protein and relative abundance of each peptide is presented as normalized values based on the total observed signal for each individual peptide. The locations of the N, D1, and D2 domains are annotated above the amino acid numbers and the position of the two GDPH motifs. *F* and *G*, recombinant FCGBP-ND2 with the natural GDPH sequence in vWD1 and vWD2 (*F*) and its GAPH-mutated counterpart (*G*) were analyzed by SDS-PAGE and stained by Coomassie Brilliant Blue or using WB with the  $\alpha$ FCGBP-3517 antiserum. The mutation inhibits cleavage, and the full molecule is detected also in reducing conditions. Schematic figures are included. FCGBP, IgGfc-binding protein; GDPH, Gly-Asp-Pro-His; vWD, von Willebrand D; WB, Western blot.

the construct was analyzed under reducing conditions, the 130-kDa protein remained intact (Fig. 3G). The increased mobility of the nonreduced proteins is caused by incomplete

unfolding of condensed protein. These experiments confirm that the autocatalytic cleavage in FCGBP is depending on the Asp/Pro sequence of the vWD domains.

### An interfragment disulfide bridge stabilizes each FCGBP cleavage site

Alignment of all vWD domains in FCGBP revealed conservation of the cysteine positions (Fig. S3A) and comparison of the vWD1 sequence of FCGBP to the same domain sequences of the vWF allowed for prediction of the pairing cysteines within the domains (Fig. S3B). There is structural information of vWD domains obtained from X-ray crystallography and cryo-EM studies of the vWF and MUC2, respectively (30, 31). Using the vWF structure, the FCGBP vWD1 domain structure was modeled to obtain further information of the potential Cys involved in the stabilization of the GDPH cleaved protein (Fig. 4A). The model suggested that the predicted Cys, C472, and C611 were positioned close enough to allow the formation of a disulfide bridge. No other cysteines were located in position to form additional associating bonds between the two cleavage products. The GDPH motif is predicted to form a loop between two beta sheets in the vWD1 domain and the GD and PH terminal ends were predicted to be buried within the structure, which could affect accessibility to react with other proteins.

### Covalent binding of FCGBP to other proteins

The identified FCGBP bands from mouse mucus separated on SDS-PAGE under nonreducing conditions were cut out and subjected to in-gel digestion and MS analysis. The analysis did not consistently reveal any other secreted proteins than FCGBP except for serum components (Table 3). Thus, we could not observe any other mucus protein covalently bound to FCGBP.

The Asp-anhydrides formed upon GDPH cleavage of FCGBP could potentially react with water. Given that the structural prediction suggested that the formed C terminus is oriented internally in the molecule, it is likely less accessible for binding (Fig. 4, A and B). Analysis of mass spectra identified few peptides ending in GD because of their size after digestion or potential modifications. However, MS analysis of the FCGBP-ND2 did not only show the semitryptic P(P478)H cleavage product of D1 but also the semi-lysyl endopeptidic peptide that ends with the amino acids GD. This anhydride had reacted with water, thereby rebuilding the original aspartate residue as the new C terminus (Fig. 4A, Fig. S3C). This result shows that this Asp-anhydride was not cross-linked to other molecules, indicating that Fcgbp might not bind other proteins this way.

To further investigate the relation between MUC2 and FCGBP, we analyzed the relative amount of mucus that was soluble and insoluble in guanidinium chloride (GuHCl), by in-solution digestion and MS analysis. The Muc2 mucin forms insoluble complexes during biosynthesis (32) and in guanidinium after storage (33). Thus, Muc2 was predominantly found in the insoluble mucus fraction (Fig. 4B). On the contrary, Fcgbp was three times more abundant in the soluble mucus fraction than the insoluble mucus with significant difference compared with Muc2 that showed more variation (Fig. 4B). As most of the Muc2 was found in the insoluble fraction, and as

unbound Fcgbp could be removed by reduction (17), the amount of Fcgbp bound to Muc2 was further investigated by in-gel digestion of Muc2 bands separated by agarose-PAGE (Fig. 4C). Fcgbp only accounted for a small (<1%) fraction of the combined intensities of these proteins in the bands (Fig. 4D). Thus, Fcgbp is mostly found as a soluble protein not bound to Muc2.

### Noncovalent binding to IgG

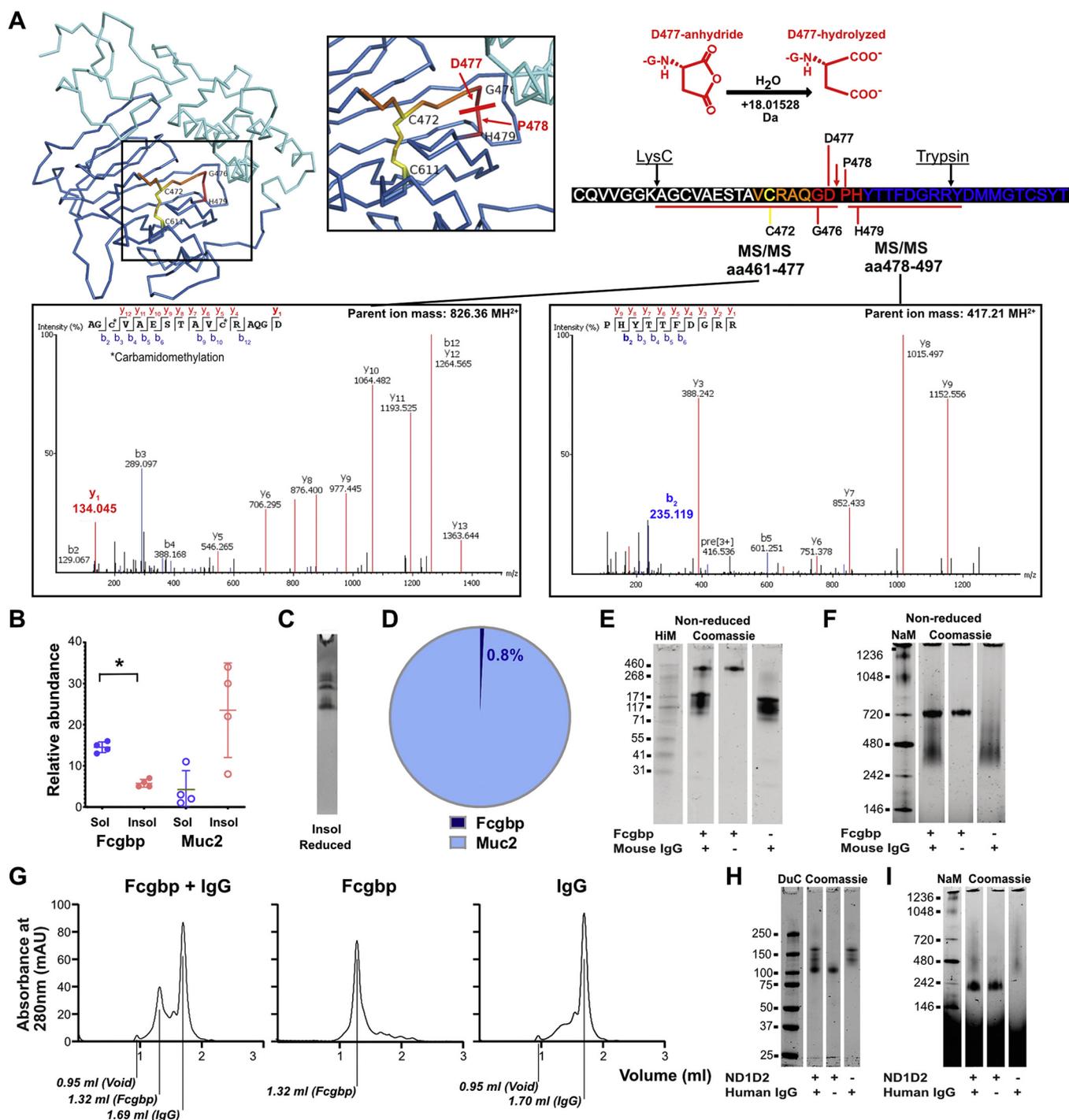
FCGBP was originally identified and named to be a protein binding to the Fc portion of all IgG subclasses (25, 26). Immunoglobulins were also detected in the nonreduced Fcgbp band from mucus (Table 3), but additional serum proteins were identified, indicating contamination with these abundant proteins (34). To further analyze this, IgG purified from mouse serum was incubated to allow binding *in vitro* with purified mouse recombinant full-length Fcgbp. Separation by SDS-PAGE under nonreducing conditions did not indicate any binding (Fig. 4E) nor did separation on gel under native conditions (Fig. 4F). This lack of interaction was further confirmed by separating the preincubated Fcgbp and IgG sample by size-exclusion chromatography (Fig. 4G). As the original analysis of IgG binding was performed on human samples and the only domain substantially differing between species is the N-terminal domain, we also analyzed a human recombinant protein containing the N-terminal domain and two adjacent vWD domains. Neither SDS-PAGE (Fig. 4H) nor native electrophoresis (Fig. 4I) showed any binding of the human FCGBP-ND2 construct to human IgG after incubation.

### Discussion

Our results show that both in mouse and human mucus, FCGBP is cleaved at all its GDPH sequences present in all but two of the vWD domains, yet held together covalently through a single cysteine bond within each vWD domain. FCGBP is expressed along the full length of the crypt in human sigmoid and murine distal colon and is secreted as an intact associated protein. The function of FCGBP is still poorly understood, but the many vWD domains of FCGB with GDPH sequences indicate interactions and previous studies have suggested it to act as a mucus cross-linker or to sequester IgG in the mucus (17, 25, 26).

The D–P bond is sensitive to acidic conditions and could potentially be cleaved during preparations, but as we observe cleavage at neutral pH without boiling and that we also observe noncleaved Fcgbp in lysates argue for a natural mechanism. Although enzymatic processing has been proposed for GDPH cleavage of some proteins (9, 10), most data support an autocatalytic mechanism (7, 8, 11–15). The GDPH cleavage of Fcgbp occurs early during biosynthesis in the endoplasmic reticulum as seen for some other GDPH-cleaved proteins (7, 9, 15). The GDPH cleavage product will generate a reactive anhydride in the new C terminus that could react with other molecules forming covalent ester or amide bonds (15, 19, 20). Such interactions would withstand SDS-PAGE separation, but no protein attached to FCGBP was identified by MS

## GDPH cleavage in FCGBP



**Figure 4. GDPH fragments are associated with a single disulfide bond, but most FCGBP is not bound to other molecules.** *A*, a model of FCGBP vWD1 structure based on the vWD structures of the vWF and MUC2. The GDPH motif is predicted to form a loop between two beta sheets. Cleavage between D477 and P478 results in the cleavage products only being covalently stabilized through one disulfide bond between C472 and C611. MS/MS spectrums for the semidigested GD and PH cleavage products from the vWD1 of recombinant FCGBP-ND2 are shown. The **P(478)HYTTFDGR** peptide is digested with trypsin, whereas the **KAGCVAESTAVCRAQGD(477)** is digested with LysC. The y-series of the MS2 spectrum suggests that the C-terminal anhydride has been hydrolyzed as indicated by the illustration on the right panel. *B*, label-free quantification (LFQ) of Muc2 and Fcgbp in the soluble or insoluble fraction after GuHCl extraction and in solution (FASP) preparation and MS/MS analysis. Fcgbp is significantly different in the two fractions, \* $p < 0.05$ ,  $n = 4$ . *C*, agarose-PAGE stained by Alcian blue detecting Muc2 bands of the GuHCl-extracted insoluble fraction. *D*, the percentage of the total Muc2 and Fcgbp combined intensities of in-gel digested and MS analyzed bands detected in panel *C*. *E*, purified mouse IgG and recombinant Fcgbp were preincubated and analyzed by Coomassie Brilliant Blue-stained SDS-PAGE in nonreducing conditions. Proteins loaded individually as indicated were used as controls. *F*, Coomassie Brilliant Blue-stained native PAGE with mouse IgG and Fcgbp preincubated or loaded individually. *G*, size-exclusion chromatography elution diagrams where recombinant Fcgbp and mouse IgG had been either mixed and injected or run separately. Elution volumes for the proteins are marked. *H*, human FCGBP-ND2 incubated with human IgG analyzed by nonreduced SDS-PAGE compared with the proteins loaded individually. *I*, samples as in panel *H* analyzed by native PAGE. FASP, filter-aided sample preparation; FCGBP, IgGfC-binding protein; GDPH, Gly-Asp-Pro-His; GuHCl, guanidinium chloride; IgG, immunoglobulin G; MUC2, mucin-2; vWD, von Willebrand D.

**Table 3**  
Secreted proteins in nonreduced <460-kDa Fcgbp band analyzed with MS

Protein	Total coverage (%)	Total area sample 1	Total area sample 2	Total area sample 3	Total peptides	Unique peptides	Average mass (Da)	UniProt accession	Source
Fcgbp	62	4.27E8	2.63E8	1.61E9	238	202	275,239	E9Q9C6	Mucus
Pregnancy zone protein	28	1.28E6	2.22E7	1.66E8	40	39	165,852	Q61838	Serum
Serum albumin	21	3.34E5	2.3E7	1.37E7	13	13	68,693	P07724	Serum
Immunoglobulin heavy constant alpha (fragment)	30	6.31E6	2.06E6	3.76E6	7	7	42,078	A0A0A6YXW	Serum
Beta-globin	22	5.21E5	7.52E5	1.23E8	3	3	15,748	A8DUK4	Serum

analysis. The Asp-anhydride can react with different nucleophiles, such as hydroxyl groups, amines, and also water (21). Structural modeling of the FCGBP vWD1 domain suggested that the Asp-anhydride should be well hidden inside the folded protein. This would rather suggest that Fcgbp uses a water molecule as the reaction partner and becomes hydrolyzed, possibly important to limit formation of large complexes during early biosynthesis.

The role of FCGBP as a mucus cross-linker binding MUC2 was addressed by GuHCl extraction of the mucus. The majority of the Fcgbp was then found in the soluble fraction but the majority of Muc2, with only small amounts of Fcgbp, were found in the insoluble fraction. When insoluble mucus was separated by agarose-PAGE, the majority of Fcgbp was not detected within the Muc2 band. Fcgbp was previously observed by us to be covalently bound to Muc2 (17). If such a complex really exists, it is only a very minor amount of Fcgbp that is bound to Muc2 in mucus. This does of course not exclude other noncovalent interactions to Muc2 or other mucus proteins. Instead, it may suggest transient interactions as opposed to the crosslinks found in the highly insoluble Muc2 mucin framework (23). Fcgbp could also, as indicated by our results, be able to form larger complexes under native conditions.

IgG and Fcgbp were, in contrast to previous suggestions, not observed to interact covalently or noncovalently (16, 25, 27). Many of the early studies used fixed tissue to determine IgG binding. Binding was observed in tissues treated with denaturing fixatives, but not if the structure was preserved by cross-linking fixatives. Both human and mouse IgG was observed to bind to human colonic tissue, but not to some other species tested (25). The IgG Fc binding was in previous work shown to include several of the vWD domains with a strong reactivity to vWD5. The N-terminal domain did not bind IgG but was indicated to have a general effect on the protein function (16). We used native purified full-length murine Fcgbp to study the IgG–FCGBP interaction and did not observe any binding to mouse IgG. The great homology between species would argue for shared functions between them. The extended length of the human FCGBP is caused by a repeat of the central triplets of the vWD assemblies, and thus, all the different types of domains are present in the mouse sequence except for the N-terminal domain. Binding to the human FCGBP N-terminal domain was tested using recombinant human FCGBP-ND2. No binding to human IgG was observed in neither native nor denaturing conditions. The evolutionary early appearance of FCGBP molecules would also not favor functions related to more recent development as the adaptive immune system (6). Binding to IgG is not observed using native proteins and is thus likely to be weak or indirect if occurring.

The GDPH-cleaved fragments are, according to our model based on the known vWD structure, associated with one disulfide bridge in each domain, similar to other GDPH-containing vWD domain proteins (11, 35), and are thus a sensitive part of the molecule. These bonds should be mechanically weak as each one is predicted to only be further stabilized by a single, fairly short beta sheet. This would reduce

## GDPH cleavage in FCGBP

the tensile strength along the axis of each molecule to a single covalent bond in each domain. We speculate that this could be of functional importance for FCGBP as shear forces are of great importance for the vWF (36). The disulfide bond of the cleaved FCGBP could also be susceptible to changes in redox potential, something that could be linked to bacterial effects (37, 38). One of the Cys in the fragment-linking disulfide bond is found within a thioredoxin motif (CGLC), which is present in each vWD domain of FCGBP as well as in mucins, but no function for this motif has been attributed yet.

In conclusion, FCGBP is fully cleaved at all its GDPH sequences but secreted as a fully disulfide bond, stabilized molecule that do not seem to bind other proteins. The great abundance could however indicate a function related to mucus gel formation and the single disulfide bond located in a potential thioredoxin motif could mediate controlled dissociation or rearrangement of the molecule depending on environmental factors. Further studies are needed to understand the functional role of this mucus protein.

## Experimental procedures

### Human samples

Sigmoid colon biopsies were obtained from patients referred for colonoscopy at Sahlgrenska University Hospital, Gothenburg, Sweden, in compliance with the Declaration of Helsinki and approved by the human Research Ethical Committee in Gothenburg, Sweden (040-08, 136-12, and 2020-03196). All included patients were adults (above age 18 years) and gave written informed consent. Patients with normal intestinal macroscopy were included.

### Animals

*C57BL/6N* (WT, in house breed), *Fcgbptm1b*(EUCOMM) *Wtsi/Cnrm* (EM05780) crossed to *C57BL/6N* background (*Fcgbp*<sup>-/-</sup>) and *Muc2*<sup>-/-</sup> (39) mice were kept at 21 to 22 °C with 12-h day/night cycles. The experiments involved gathering samples from mice with an age span of 8 to 20 weeks. Mice were euthanized through sedation with isoflurane, followed by cervical dislocation. Animal experiments were approved by the Ethical Committee on Animal Experiments in Gothenburg, Sweden (number: 280-2012, 73-2015, and 2285-2019).

### Recombinant FCGBP/*Fcgbp* constructs

The vector pCMV-*Fcgbp* (amino acids 1–2583, NP\_001116075.1)-Myc-DDK (NM\_001122603, OriGene) was used to express full-length mouse *Fcgbp*. The plasmid pCMV-FCGBP (amino acids 1–5405, NP\_003881.2)-Myc-DDK (NM\_003890, OriGene) was used as a template for all the FCGBP fusion proteins. FCGBP-ND2 (amino acids 24–1250, NP\_003881.2) was generated through PCR (Pfu turbo DNA polymerase, Agilent) using the forward primer *GGGAAGCTTCAGTCGACCTCAAGAACTGG* and the reverse primer *GGGTCTAGAGCTAGAGCCCACGGCGACACAGC* (Eurofins Genomics). The insert was introduced into the XbaI- and HindII-cleaved pSecTag(+)-Myc-His C vector

(Invitrogen, Thermo Fisher Scientific). The Gly-Ala-Pro-His (GAPH) double mutant was produced in two steps. First, the vWD1 GDPH to GAPH was mutated using the primers *GCGCCAGGGCGCCCCATTACAC* and *GTGTAATGGGGCGCCCTGGGCGC*, and second, the primers *CCAGGGGTCCGGGGCCCCACACTATGTGA* and *TCACATAGTGTGGGGCCCCGGACCCCTGG* (Eurofins Genomics) were used for the GDPH to GAPH mutation in vWD2. The Gibson Assembly Master Mix kit (New England Biomedicine) (40) was used to assemble the plasmid containing the FCGBP N-terminal (amino acids 1–470, NP\_003881.2) into an EcoRV linearized pcDNA3.1(+)-Myc-His B vector (Invitrogen, Thermo Fisher Scientific). The forward primer *ATCCACTAGTCCAGTGTGGTGGAATTCTGCAGATATGGGTGCCCTATGGAGCTGGTGATACTCTGGGCTGGAGCAACCCTCTGTGGGGATTGACCCAGGAGGCTTCAGTGGACCTCAAGAACACTGG* and the reverse primer *GACTCGAGCGGCCGCCACTGTGCTGGATAGCGG TGGACTCCGCCACACCCGGCCTTCCCGCCTACC* were used. The FCGBP-D13 (amino acids 5225–5405, NP\_003881.2) was generated in the same way using the forward PCR primer *AGGCGCGCCGTACGATGTGTCCTGTCTGTGGGTGCCAACCTC* and the reverse primer *TTTTTGTTCGGGCCCCCATAACATGGGGAGAAGTCCTGCGC* and was inserted into the Hind III and Apa I linearized pSecTag(+)-Myc-His A vector (Invitrogen).

### Recombinant FCGBP protein expression and purification

Plasmids were transformed into competent *Escherichia coli* XL1-Blue (Agilent), and the DNA was purified using Qiagen plasmid Mini or Maxi kits (QIAGEN) according to manufacturer's instructions. The truncated proteins were transiently expressed using FectoPRO (Polyplus) in suspension-growing CHO-S cells and prepared from 300 ml (FCGBP-ND2 GDPH and FCGBP-ND2 GAPH mutant) and 1000 ml (FCGBP-N and D13) of FectoCHO (Polyplus) cell medium. The buffer was exchanged to PBS through tangential flow filtration. The FCGBP-ND2 GDPH/GAPH in PBS with 500 mM NaCl was loaded onto a 1-ml cobalt-loaded HiTrap Chelating FF affinity column (GE Healthcare) previously equilibrated with PBS. Three elution steps were performed using 20, 200, and 500 mM of imidazole (50 mM NaHPO<sub>4</sub>, pH 7.4, 500 mM NaCl), respectively. FCGBP-ND2-containing fractions were pooled and dialyzed (20 mM Hepes, pH 7.4, 150 mM NaCl) overnight using a 6- to 8-kDa cut-off dialysis membrane tubing (Spectra-Por). The FCGBP-N and FCGBP-D13 were instead loaded onto 1-ml nickel-based Hi-trap affinity columns (GE Healthcare) using 20, 50, and 500 mM imidazole for stepwise elution. Buffer exchange of the protein fractions to sterile PBS was performed using a Vivaspin 20 with 10-kDa cut-off (Sartorius).

Full-length mouse *Fcgbp* was expressed in CHO-S by transient transfections of the *Fcgbp*-Myc-DDK plasmid using the FectoPRO transfection reagent (Polyplus). The medium was dialyzed to 50 mM Hepes, pH 7.4, and 50 mM NaCl using a 6- to 8-kDa cut-off membrane (Spectra-Por). The material

was loaded at 4 °C into a gravity column packed with 3 ml EZview Red ANTI-FLAG M2 affinity gel (Sigma). The column was washed with PBS, and the protein was eluted with 100 µg/ml 3X FLAG peptide (Sigma) in PBS. Further PBS wash and elution using 100 mM glycine, pH 3.5, were done, followed by the neutralization of the samples during collection using 500 mM Tris/HCl, pH 7.4, and 1.5 M NaCl. The FLAG eluate and subsequent PBS washes were pooled and concentrated using a 30-kDa cut-off Vivaspin 20 column (Sartorius). A final size-exclusion purification was done using a Superose 6 Increase 10/300 column (GE Healthcare) equilibrated in 10 mM Tris/HCl, pH 7.4, with 100 mM NaCl using an ÄKTA purifier system (GE Healthcare) at room temperature (RT). Alternatively, Fcgbp was purified using a MONO Q 5/50 or HiPrep Q XL (GE Healthcare) coupled with the ÄKTA system before size-exclusion purification using a Superose 6 column. The eluate was concentrated, buffer exchanged to 25 mM HEPES, pH 7.4, with 100 to 150 mM NaCl, and stored at -80 °C.

To generate a stable clone of the full-length mouse Fcgbp in CHO-K1 cells, 70% confluent cells grown in Iscove's Modified Dulbecco's Medium with 10% fetal calf serum (Gibco) were transfected using the Lipofectamine 2000 kit (Thermo Fisher). Fcgbp-expressing cells were selected for 4 weeks by addition of Geneticin G418 (800 µg/ml, Gibco) 48 h after transfection. Clones were expanded and screened by WB. Lysates from Fcgbp-expressing cells were prepared in lysis buffer (50 mM Tris/HCl pH 8, 150 mM NaCl, 1% Triton X-100) with 1× complete EDTA-free protease inhibitor cocktail (Roche) and treated with 0.1 U EndoH (Roche) in 190 mM Tris/HCl pH 7.4 with azide overnight at 37 °C.

#### Phylogenetic mapping and signal peptide prediction

The protein sequences of all the vWD assemblies in mouse (NP\_001116075\_1) and human FCGBP (NP\_003881.2) were aligned with the vWD domains of the human vWF (41). This allowed for extrapolation of the domain borders based on the distribution of the Cys. The web-based phylogenetic analysis tool (<https://www.phylogeny.fr/>) was used to generate a tree (42).

The signal peptides of the human and mouse FCGBP sequences were predicted using SignalP v. 5.0 (43).

#### FCGBP/Fcgbp antibodies

Purified FCGBP-D13 (antigen for αFCGBP-D13/αD13) and FCGBP-N (antigen for αFCGBP-N/αN) were used to immunize rabbits. Immunization was performed six times with 4-week intervals (Agrisera AB). The specificity of the antisera was tested using immunohistochemistry and WB. Interspecies reactivity was predicted. Because the αFCGBP-D13 antigen showed a 70% sequence homology with the mouse vWD7, it was predicted to detect the C-terminal GDPH cleavage product in both species. The N-terminal antigen from humans showed a 79% homology with the much shorter N-terminal part of mouse Fcgbp (amino acid 26–59). The αFCGBP-3517 polyclonal antiserum (HPA003517, Atlas Antibodies) targets mainly the vWD5, vWD8, and vWD11 domains, which in

mouse best match a sequence in vWD5 with some minor homology with the rest of the vWD domains. The αFCGBP-3564 (HPA003564, Atlas Antibodies) specifically targets the N-terminal domain of human FCGBP with no predicted cross-reactivity with vWD domains nor with the truncated mouse N terminus. The MUC2C3 (αMUC2C3/αC3) antiserum (1) targeting the GDPH peptide PHYVTFDGLYYSYQGNC of MUC2 was a good candidate for detecting human and mouse Fcgbp fragments. For humans, the best sequence match was found in the repeated domains vWD5, vWD8, and vWD11, whereas the vWD3 was the best match for mouse Fcgbp. Table S1 shows an overview of the antibodies.

#### Immunohistochemistry

Human colonic sigmoid biopsies and mouse distal colons were immediately fixed in methanol-Carnoy's fixative for 48 h with 100% methanol used for long-term storage. The samples were paraffin-embedded and cut into 4-µm section (1, 44). Slides were deparaffinized in Xylene substitute (Sigma) and stepwise hydrated. The mouse sections were subjected to antigen retrieval in 10 mM citric acid buffer, pH 6, at 100 °C for 30 min. The samples were blocked using 2% BSA in PBS, and primary antibodies were incubated overnight at 4 °C. For mouse tissue, the αMUC2C3, αFCGBP-D13, and αFCGBP-N sera were used at 1:2000 dilutions and at 1:4000 for human samples. FCGBP pre-serums (before immunization) for αFCGBP-N and αFCGBP-D13 were used as negative controls. Goat anti-rabbit-Alexa 555 (Thermo Fisher Scientific) diluted 1:2000 was used as the secondary antibody, and DNA was counterstained with Hoechst-34580 (1 µg/µl, Thermo Fisher Scientific). The sections were mounted in ProLong Gold antifade (Thermo Fisher Scientific). Images were acquired using a Nikon Eclipse E1000 fluorescence microscope with a 20× 0.50 NA Plan Fluor DIC objective (Nikon). The 64-bit NIS-Element D (version 4.20.01, Nikon) and Imaris Viewer (version 9.5.1, Imaris) were used for image processing.

#### Extraction of human and mouse colonic cell lysate and secreted mucus

For extraction of human colonic mucus, the tissue was flushed and stored in Krebs transport buffer before mounting in an explant chamber system (45). The apical surface was submerged in 150 µl of Krebs-mannitol buffer. After 1 h of mucus growth, the buffer containing the scraped mucus was extracted and stored at -80 °C. For extraction of mouse mucus, the dissected and opened distal to mid-colon was mounted with the epithelium facing upward on a silicon-coated plate. The tissue was overlaid with 150 to 200 µl of PBS and the mucus gently scraped off. Human and mouse colonic biopsies were lysed in 400 µl of the lysis buffer (50 mM Tris/HCl pH 8, 150 mM NaCl, 1% Triton X-100) with 1× complete EDTA-free protease inhibitor cocktail (Roche) and homogenized using an Ultra Turrax IKA (Werke), followed by incubation at RT for 20 min. Samples were centrifuged at 9000g for 15 min at 4 °C. All samples were stored at -20 °C.

## GDPH cleavage in FCGBP

### Electrophoresis and WB

Samples for SDS-PAGE were prepared in the sample buffer (50 mM Tris/HCl, pH 6.8, 2% SDS, and 10% (v/v) glycerol and bromophenol blue) and incubated at 37 °C for 30 min without (nonreduced) or with addition of 100 mM DTT (reduced) followed by incubation at 95 °C for 5 min before loading, unless otherwise stated. Samples were separated on gels (Mini-PROTEAN TGX 4–15, 20% Bio-Rad and 4, 8 or 10% polyacrylamide) and size estimation was performed using the Precision Plus Protein Dual color standard (Bio-Rad) or the HiMark pre/unstained ladder (Thermo Fisher Scientific). For sample comparison (*Fcgbp*<sup>-/-</sup> to WT), concentrations were normalized based on protein amount measurements using the Pierce BCA assay kit (Thermo Fisher Scientific). For optimal gel separation and coverage by MS, His-purified FCGBP-ND2 was denatured in 0.05% (w/v) SDS and reduced by the addition of DTT (10 mM) for 1 h at 60 °C, followed by alkylation for 15 min at RT using iodoacetamide (IAA) (20 mM). Primary amines were converted to dimethylamines by adding 1- $\mu$ l 4% formaldehyde and 1- $\mu$ l 0.6 M sodium cyanoborohydride to the 25  $\mu$ l sample and incubating for 4 h at 37 °C. Proteins were precipitated by the addition of one volume of 20% trichloroacetic acid and eight volumes of ice-cold acetone and left on ice for 1 h followed by centrifugation for 15 min at 14,000g. Pellets were resolved in the sample buffer and analyzed by SDS-PAGE on an 8% gel and stained with Imperial protein stain (Thermo Fisher Scientific).

Native PAGE (4–16% Bis-Tris gel) was used to analyze samples prepared in Native PAGE sample buffer (Thermo Fisher Scientific) without a sample additive. Native PAGE gels were run at 4 °C first with the dark blue cathode buffer for 60 min at 150 V and then switched to light blue cathode buffer and run at 250 V for 90 min. Gels were stained with Imperial Protein stain (Thermo Fisher Scientific) overnight at RT and later washed in MilliQ water.

Separation of the large Muc2 bands in mucus was performed by agarose-PAGE as previously described (46).

Proteins were semidry-blotted to Immobilon-P polyvinylidene difluoride (PVDF), or PVDF-FL membrane (Millipore) in transfer buffer (48 mM Tris HCl, 21 mM glycine, and 1.3 mM SDS). Reduced gels were blotted for 60 min at 2.5 mA/cm<sup>2</sup>, while nonreduced and native PAGE gels were blotted for 80 min using the same parameters.

For PVDF-FL membranes stained with fluorescent antibodies, PBS with 5% BSA was used for both blocking and diluting the primary antibody ( $\alpha$ FCGBP-D13 and  $\alpha$ FCGBP-N 1:5000). As the secondary antibody, goat anti-rabbit IgG (H + L) Alexa Fluor 790 (Thermo Fischer Scientific) was added 1:20,000 in PBS-T (0.01% Tween 20) with 5% BSA and 0.02% SDS and incubated in dark at RT for 1 h. Gels were imaged using an Odyssey CLx imager (Li-Cor Biotechnology).

The membranes developed with alkaline phosphatase were blocked in PBS-T with 5% milk. The primary antibodies used were  $\alpha$ FCGBP-3517 (1:200, HPA003517, Atlas Antibodies),  $\alpha$ FCGBP-3564 (1:200, HPA003564, Atlas Antibodies),  $\alpha$ MUC2C3 (1:2000) (1),  $\alpha$ FCGBP-D13 (1:2000),  $\alpha$ FCGBP-N

(1:2000), or M2 $\alpha$ Flag (1:2000, Sigma). An IgG (H + L) alkaline phosphatase (1:1000–2000, SouthernBiotech) was used as the secondary antibody and blots were developed using premixed bromo-4-chloro-3-indolyl phosphate/nitroblue tetrazolium (SouthernBiotech).

Full protein content, based on staining by Imperial Protein stain (Thermo Fisher Scientific) of gels or Revert 700 total protein stain (Li-Cor Biotechnology) of membranes was used as loading controls.

### GuHCl extraction of mucus

Scraped mucus from the mouse distal colon was directly transferred to 200- $\mu$ l GuHCl buffer (10 mM sodium phosphate, 6 M GuHCl, 5 mM N-ethyl-maleimide, 5 mM EDTA, and 0.5 mM PMSE, pH 6.5). Proteins were extracted for 12 h at 6 °C under gentle shaking. Insoluble mucins were precipitated by centrifugation at 16,000g followed by a wash in GuHCl buffer. Supernatants (soluble fractions) for each sample were combined for MS analysis or separated by agarose-PAGE (46).

The insoluble pellet sample for electrophoresis was brought into solution by the addition of 100 mM Tris, 5 mM EDTA, pH 8.0, and freshly added 10 mM DTT under gentle shaking for 1.5 h at 37 °C. The cysteinyl groups were alkylated by the addition of 25 mM IAA. Samples were dialyzed against water and concentrated in a vacuum centrifuge before separation on composite agarose-PAGE.

### MS sample preparation

Protein samples for MS analysis were either digested in-solution or from electrophoresis gel bands. Mucus samples from mouse and human colons were prepared as previously described (2, 3). Lysates, GuHCl-extracted samples, and recombinantly expressed proteins were digested in-solution using a modified filter-aided sample preparation protocol (2, 3, 47). Samples were reduced with 10 mM DTT under gentle shaking for 2 h at 37 °C, transferred to 10-kDa cut-off filter units (Nanosep, Pall Lifescience) and centrifuged at 10,000g. The samples were alkylated by addition of 50 mM IAA in 100 mM Tris (pH 8.5) on the filter. After washing the filter with 50 mM ammonium bicarbonate, pH 8.0, LysC (10 ng/ $\mu$ l, Wako, cleaving at Lys) or trypsin (10 ng/ $\mu$ l, Promega, cleaving at Lys and Arg) were used for overnight digestion at 37 °C. Peptides were eluted in 50 mM ammonium bicarbonate or 500 mM NaCl, acidified, and purified through C18 Stage Tips (48). The peptides were vacuum dried and resolved in 0.1 to 0.2% formic acid before MS analysis.

Selected gel bands were dissected and prepared through in-gel digestion (23, 49) (modified from (50)). Gel pieces were destained in 50% acetonitrile and reduced overnight at 37 °C or 1 h at 56 °C in 10 mM DTT. Samples were alkylated with 25 mM IAA for 20 min at RT protected from light. Between each step, the gel pieces were washed in acetonitrile. Digestion was performed using trypsin or LysC in 50 mM ammonium bicarbonate, pH 8.0, or AspN (10 ng/ $\mu$ l, Promega, cleaving at Asp but to some extent also at Glu) in 25 mM Tris/HCl, pH 7.8, with overnight incubation at 37 °C. Extracted peptides were

vacuum-dried before C18 stage-tip cleaning (48). The peptides were resolved in 0.1 to 0.2% formic acid for MS analysis.

### MS and data analysis

The protein composition of the samples was studied by nanoLC-ESI-MS/MS as described before (2, 3, 51). Samples were analyzed with an LTQ-Orbitrap, an LTQ-Orbitrap XL (2, 3), or Q-Exactive mass spectrometers (2, 23) (Thermo Fisher Scientific). Peptides were separated with an in-house column (150 mm × 0.075 mm inner diameter, New Objective) packed with ReproSil-Pur C18-AQ 3 μm particles (Dr Maisch), using a 5 to 35% acetonitrile gradient. Full mass spectra were acquired over a mass range of minimum 200 m/z and maximum 2000 m/z, with a resolution of at least 60,000. The 6 to 15 most intense peaks with a charge state ≥2 were fragmented with normalized collision energy of up to 30%, and tandem MS was acquired at a resolution of 17,500.

The raw data were converted using ProteoWizard 3.0.18146 and analyzed by Mascot (version 2.6.0., Matrix Science) or PEAKS 8.5 (2017 Bioinformatics Solutions Inc). Label-free quantification was performed using MaxQuant (version 1.5.7.4) (52). In general, enzyme cleavage was set as semi-specific, and 2 to 3 missed cleavages were allowed. Fixed modifications were carbamidomethylation of cysteines, and variable modifications were oxidation of methionines and acetylation of the protein N terminus in Mascot and the 313 built-in modifications in PEAKS. Searches were made against specific protein sequences of the constructs, selected UniProt entries as mouse Muc2 and mouse or human FCGBP (Q80Z19, E9Q9C6, and Q9Y6R7, respectively) or full databases (mouse: Uniprot10090, 84,123 sequences; human: Uniprot9606, 93,555 sequences; in-house database available online (<http://www.medkem.gu.se/mucinbiology/databases/index.html>): MUCextra (18 sequences)). Mass tolerance did not exceed 10 ppm for precursor ions and 1 Da for fragments. For in-gel digested bands, a Mascot ion score cut-off of 25 was used. Peptides were exported and sorted based on which putative GDPH-cleavage fragment they originate from. Heat maps were generated using GraphPad Prism 8, showing the percentage of total identified tryptic Fcgbp peptides from each putative GDPH cleavage product. N-terminally labeled FCGBP-ND2 peptide identifications were accepted based on the Mascot ion score (≥20) with searches including additional fixed modifications of dimethylation (K, N-terminal) and methylation (P). Relative abundance of all identified FCGBP-ND2 peptides was determined based on the extracted ion chromatogram using MS1 filtering in Skyline (53). The signal for each peptide was summed, and the percentage in the individual bands was calculated to determine which band was containing the different autocatalytic fragments.

### FCGBP-D1 model prediction

A model for the D1 assembly of human FCGBP protein (amino acids 469–800, NP\_003881.2) was obtained using the 3D modeling prediction program MODELLER (54), with the D'D3 structure of the vWF (PDB ID: 6N29) as the template (30). The structural prediction shows the putative location of

the GDPH motif in relation to the cysteine bonds that stabilize the vWD domain after cleavage.

### IgG-binding assays

Purified recombinant Fcgbp- (1.5 μg) and mouse IgG purified from serum (1.5 μg, cat# I5381, Sigma) or human FCGBP-ND2 (1.5 μg) and human IgG also purified from serum (1.5 μg, cat#I4506, Sigma) were incubated together for 20 min at RT in 25 mM Hepes, pH 7.4, with 150 mM NaCl before electrophoresis. Analytical gel-filtration was performed using a Superose 6 increase 3.2/300 mm column (GE Healthcare) coupled to an Ettan LC system (GE Healthcare) and using the software Unicorn 5.01 (Build 318, Amersham Biosciences 2004). Purified Fcgbp and IgG were either run separately or injected together in a 2:1 (IgG:Fcgbp) concentration (mg/ml) ratio. The fractions were analyzed by SDS-PAGE.

### Statistics

Data are presented as individual values with the median and differences calculated using nonparametric testing (Mann-Whitney U). The significance level was  $p < 0.05$ , and  $p$  was two sided. The relative amount of the two components in a gel band was calculated as percentages of their total intensities. Heat maps were generated based on peptide coverage per domain and presented as the percentage of all coverage. Statistics were calculated using GraphPad Prism (version 9.0.2).

### Data availability

Public raw mass spectrometry data from human mucus were obtained from the ProteomeXchange repository with the dataset identifier PXD012632 (2). The mass spectrometry proteomics data generated have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) with the dataset identifiers PXD025496 (mouse colon whole mucus) and PXD025493 (mouse colon mucus extractions).

Additional MS data are available as Ehrencrona, Erik (2021), "FCGBP, the IgGFc-binding protein, is secreted with all GDPH sequences cleaved, but maintained by inter-fragment disulfide bonds", Mendeley Data, V1, <https://doi.org/10.17632/bkzwmd35jt.1> <https://doi.org/10.17632/bkzwmd35>.

**Supporting information**—This article contains [supporting information](#).

**Acknowledgments**—We thank the Mammalian Protein Expression Core Facility and the Centre for Cellular Imaging at the Sahlgrenska Academy for their assistance. We are thankful for all technical assistance and help from Frida Svensson and Joakim Bergström. This work was supported by the Swedish Research Council (2015-03047, 2017-00958, 2019-01134), the Swedish Cancer Foundation (CAN2016/487, 2017/360), the Knut and Alice Wallenberg Foundation (2017-0028), National Institute of Allergy and Infectious Diseases (U01AI095473), European Research Council (ERC) (694181), IngaBritt and Arne Lundberg Foundation (2015-070,

## GDPH cleavage in FCGBP

2018-0117), Sahlgrenska University Hospital (ALFGBG-724681, ALFGBG-440741), and the Sahlgrenska Academy.

**Author contributions**—E. E., G. C. H., and M. E. V. J. conceptualization; E. E. and M. E. V. J. resources; E. E., S. v. d. P., P. G., and M. E. V. J. data curation; E. E., S. v. d. P., P. G., C. V. R., A. M. R.-P., M.-J. G.-B., M. B., and G. C. H. formal analysis; E. E., G. C. H., and M. E. V. J. supervision; E. E., G. C. H., and M. E. V. J. funding acquisition; E. E., S. v. d. P., P. G., and C. V. R. validation; E. E. investigation; E. E., S. v. d. P., P. G., and C. V. R. visualization; E. E., S. v. d. P., P. G., C. V. R., A. M. R.-P., M.-J. G.-B., S. T.-M., and M. B. methodology; E. E. and M. E. V. J. writing—original draft; E. E. and M. E. V. J. project administration; E. E., S. v. d. P., P. G., C. V. R., A. M. R.-P., M.-J. G.-B., S. T.-M., M. B., G. C. H., and M. E. V. J. writing—review and editing.

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

**Abbreviations**—The abbreviations used are: CHO, Chinese hamster ovary; EndoH, endoglycosidase H; FCGBP, IgGfC-binding protein (Fcgbp, mouse); GAPH, Gly-Ala-Pro-His; GDPH, Gly-Asp-Pro-His; GuHCl, guanidinium chloride; IAA, iodoacetamide; IgG, immunoglobulin G; ITH3, inter-alpha-trypsin inhibitor heavy chain 3; MUC2, mucin-2 (Muc2, mouse); MUC5AC, mucin-5AC; PVDF, polyvinylidene difluoride; vWD, von Willebrand D domain; vWF, von Willebrand factor; WB, Western blot.

### References

- Johansson, M. E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G. C. (2008) The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15064–15069
- van der Post, S., Jabbar, K. S., Birchenough, G., Arike, L., Akhtar, N., Sjøvall, H., Johansson, M. E. V., and Hansson, G. C. (2019) Structural weakening of the colonic mucus barrier is an early event in ulcerative colitis pathogenesis. *Gut* **68**, 2142–2151
- Rodriguez-Pineiro, A. M., Bergstrom, J. H., Ermund, A., Gustafsson, J. K., Schutte, A., Johansson, M. E., and Hansson, G. C. (2013) Studies of mucus in mouse stomach, small intestine, and colon. II. Gastrointestinal mucus proteome reveals Muc2 and Muc5ac accompanied by a set of core proteins. *Am. J. Physiol. Gastrointest. Liver Physiol.* **305**, G348–G356
- Nystrom, E. E. L., Birchenough, G. M. H., van der Post, S., Arike, L., Gruber, A. D., Hansson, G. C., and Johansson, M. E. V. (2018) Calcium-activated chloride channel regulator 1 (CLCA1) controls mucus expansion in colon by proteolytic activity. *EBioMedicine* **33**, 134–143
- Schutte, A., Ermund, A., Becker-Pauly, C., Johansson, M. E., Rodriguez-Pineiro, A. M., Backhed, F., Muller, S., Lottaz, D., Bond, J. S., and Hansson, G. C. (2014) Microbial-induced meprin beta cleavage in MUC2 mucin and a functional CFTR channel are required to release anchored small intestinal mucus. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 12396–12401
- Lang, T., Klasson, S., Larsson, E., Johansson, M. E., Hansson, G. C., and Samuelsson, T. (2016) Searching the evolutionary origin of epithelial mucus protein components—mucins and FCGBP. *Mol. Biol. Evol.* **33**, 1921–1936
- Lidell, M. E., and Hansson, G. C. (2006) Cleavage in the GDPH sequence of the C-terminal cysteine-rich part of the human MUC5AC mucin. *Biochem. J.* **399**, 121–129
- Lidell, M. E., Johansson, M. E. V., and Hansson, G. C. (2003) An autocatalytic cleavage in the C-terminus of the human MUC2 mucin occurs at the low pH of the late secretory pathway. *J. Biol. Chem.* **278**, 13944–13951
- Patrick, M. E., and Eglund, K. A. (2019) SUSD2 proteolytic cleavage requires the GDPH sequence and inter-fragment disulfide bonds for surface presentation of galectin-1 on breast cancer cells. *Int. J. Mol. Sci.* **20**, 3814
- Soto, P., Zhang, J., and Carraway, K. L. (2006) Enzymatic cleavage as a processing step in the maturation of Muc4/sialomucin complex. *J. Cell. Biochem.* **97**, 1267–1274
- Lockhart-Cairns, M. P., Lim, K. T. W., Zuk, A., Godwin, A. R. F., Cain, S. A., Sengle, G., and Baldock, C. (2019) Internal cleavage and synergy with twisted gastrulation enhance BMP inhibition by BMPER. *Matrix Biol.* **77**, 73–86
- Bell, C. H., Healey, E., van Erp, S., Bishop, B., Tang, C., Gilbert, R. J. C., Aricescu, A. R., Pasterkamp, R. J., and Siebold, C. (2013) Structure of the repulsive guidance molecule (RGM)-neogenin signaling hub. *Science* **341**, 77–80
- Thuveson, M., and Fries, E. (1999) Intracellular proteolytic processing of the heavy chain of rat pre-alpha-inhibitor. The COOH-terminal propeptide is required for coupling to bikunin. *J. Biol. Chem.* **274**, 6741–6746
- Schirrmeyer, J., Friedrich, L., Wenzel, M., Hoppe, M., Wolf, C., Gottfert, M., and Zehner, S. (2011) Characterization of the self-cleaving effector protein NopE1 of *Bradyrhizobium japonicum*. *J. Bacteriol.* **193**, 3733–3739
- Osicka, R., Prochazkova, K., Sulc, M., Linhartova, I., Havlicek, V., and Sebo, P. (2004) A novel “clip-and-link” activity of repeat in toxin (RTX) proteins from gram-negative pathogens. Covalent protein cross-linking by an Asp-Lys isopeptide bond upon calcium-dependent processing at an Asp-Pro bond. *J. Biol. Chem.* **279**, 24944–24956
- Harada, N., Iijima, S., Kobayashi, K., Yoshida, T., Brown, W. R., Hibi, T., Oshima, A., and Morikawa, M. (1997) Human IgGfC binding protein (FcgammaBP) in colonic epithelial cells exhibits mucin-like structure. *J. Biol. Chem.* **272**, 15232–15241
- Johansson, M. E., Thomsson, K. A., and Hansson, G. C. (2009) Proteomic analyses of the two mucus layers of the colon barrier reveal that their main component, the Muc2 mucin, is strongly bound to the Fcgbp protein. *J. Proteome Res.* **8**, 3549–3557
- Piszkiwicz, D., Landon, M., and Smith, E. L. (1970) Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations. *Biochem. Biophys. Res. Commun.* **40**, 1173–1178
- Engchild, J. J., Salvesen, G., Hefta, S. A., Thogersen, I. B., Rutherford, S., and Pizzo, S. V. (1991) Chondroitin 4-sulfate covalently cross-links the chains of the human blood protein pre-alpha-inhibitor. *J. Biol. Chem.* **266**, 747–751
- Morelle, W., Capon, C., Balduyck, M., Sautiere, P., Kouach, M., Michalski, C., Fournet, B., and Mizon, J. (1994) Chondroitin sulphate covalently cross-links the three polypeptide chains of inter-alpha-trypsin inhibitor. *Eur. J. Biochem.* **221**, 881–888
- Kaczmarczyk, A., Thuveson, M., and Fries, E. (2002) Intracellular coupling of the heavy chain of pre-alpha-inhibitor to chondroitin sulfate. *J. Biol. Chem.* **277**, 13578–13582
- Blom, A. M., Thuveson, M., and Fries, E. (1997) Intracellular coupling of bikunin and the heavy chain of rat pre-alpha-inhibitor in COS-1 cells. *Biochem. J.* **328**(Pt 1), 185–191
- Recktenwald, C. V., and Hansson, G. C. (2016) The reduction-insensitive bonds of the MUC2 mucin are isopeptide bonds. *J. Biol. Chem.* **291**, 13580–13590
- Thuveson, M., and Fries, E. (2000) The low pH in trans-Golgi triggers autocatalytic cleavage of pre-alpha-inhibitor heavy chain precursor. *J. Biol. Chem.* **275**, 30996–31000
- Kobayashi, K., Blaser, M. J., and Brown, W. R. (1989) Identification of a unique IgG Fc binding site in human intestinal epithelium. *J. Immunol.* **143**, 2567–2574
- Kobayashi, K., Hamada, Y., Blaser, M. J., and Brown, W. R. (1991) The molecular configuration and ultrastructural locations of an IgG Fc binding site in human colonic epithelium. *J. Immunol.* **146**, 68–74
- Kobayashi, K., Ogata, H., Morikawa, M., Iijima, S., Harada, N., Yoshida, T., Brown, W. R., Inoue, N., Hamada, Y., Ishii, H., Watanabe, M., and Hibi, T. (2002) Distribution and partial characterisation of IgG Fc binding protein in various mucin producing cells and body fluids. *Gut* **51**, 169–176
- Schwartz, J. L. (2014) Fcgbp - a potential viral trap in RV144. *Open AIDS J.* **8**, 21–24
- Albert, T. K., Laubinger, W., Muller, S., Hanisch, F. G., Kalinski, T., Meyer, F., and Hoffmann, W. (2010) Human intestinal TFF3 forms

- disulfide-linked heteromers with the mucus-associated FCGBP protein and is released by hydrogen sulfide. *J. Proteome Res.* **9**, 3108–3117
30. Dong, X., Leksa, N. C., Chhabra, E. S., Arndt, J. W., Lu, Q., Knockenhauer, K. E., Peters, R. T., and Springer, T. A. (2019) The von Willebrand factor D'D3 assembly and structural principles for factor VIII binding and concatemer biogenesis. *Blood* **133**, 1523–1533
  31. Javitt, G., Khmelnsky, L., Albert, L., Bigman, L. S., Elad, N., Morgenstern, D., Ilani, T., Levy, Y., Diskin, R., and Fass, D. (2020) Assembly mechanism of mucin and von Willebrand factor polymers. *Cell* **183**, 717–729.e716
  32. Axelsson, M. A. B., Asker, N., and Hansson, G. C. (1998) O-glycosylated MUC2 monomer and dimer from LS 174T cells are water-soluble, whereas larger MUC2 species formed early during biosynthesis are insoluble and contain nonreducible intermolecular bonds. *J. Biol. Chem.* **273**, 18864–18870
  33. Herrmann, A., Davies, J. R., Lindell, G., Martensson, S., Packer, N. H., Swallow, D. M., and Carlstedt, I. (1999) Studies on the “insoluble” glycoprotein complex from human colon. *J. Biol. Chem.* **274**, 15828–15836
  34. Adkins, J. N., Varnum, S. M., Auberry, K. J., Moore, R. J., Angell, N. H., Smith, R. D., Springer, D. L., and Pounds, J. G. (2002) Toward a human blood serum proteome: Analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteomics* **1**, 947–955
  35. Serpe, M., Umulis, D., Ralston, A., Chen, J., Olson, D. J., Avanesov, A., Othmer, H., O'Connor, M. B., and Blair, S. S. (2008) The BMP-binding protein Crossveinless 2 is a short-range, concentration-dependent, biphasic modulator of BMP signaling in *Drosophila*. *Dev. Cell* **14**, 940–953
  36. Springer, T. A. (2014) von Willebrand factor, Jedi knight of the bloodstream. *Blood* **124**, 1412–1425
  37. Espey, M. G. (2013) Role of oxygen gradients in shaping redox relationships between the human intestine and its microbiota. *Free Radic. Biol. Med.* **55**, 130–140
  38. Reese, A. T., Cho, E. H., Klitzman, B., Nichols, S. P., Wisniewski, N. A., Villa, M. M., Durand, H. K., Jiang, S., Midani, F. S., Nimmagadda, S. N., O'Connell, T. M., Wright, J. P., Deshusses, M. A., and David, L. A. (2018) Antibiotic-induced changes in the microbiota disrupt redox dynamics in the gut. *Elife* **7**, e35987
  39. Velcich, A., Yang, W. C., Heyer, J., Fragale, A., Nicholas, C., Viani, S., Kucherlapati, R., Lipkin, M., Yang, K., and Augenlicht, L. (2002) Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* **295**, 1726–1729
  40. Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345
  41. Zhou, Y. F., Eng, E. T., Zhu, J., Lu, C., Walz, T., and Springer, T. A. (2012) Sequence and structure relationships within von Willebrand factor. *Blood* **120**, 449–458
  42. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., Lefort, V., Lescot, M., Claverie, J. M., and Gascuel, O. (2008) Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469
  43. Almagro Armenteros, J. J., Tsirigos, K. D., Sonderby, C. K., Petersen, T. N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423
  44. Erickson, N. A., Nystrom, E. E., Mundhenk, L., Arike, L., Glaubien, R., Heimesaat, M. M., Fischer, A., Bereswill, S., Birchenough, G. M., Gruber, A. D., and Johansson, M. E. (2015) The goblet cell protein Clca1 (alias mClca3 or Gob-5) is not required for intestinal mucus synthesis, structure and barrier function in naive or DSS-challenged mice. *PLoS One* **10**, e0131991
  45. Gustafsson, J. K., Ermund, A., Johansson, M. E., Schutte, A., Hansson, G. C., and Sjoval, H. (2012) An ex vivo method for studying mucus formation, properties, and thickness in human colonic biopsies and mouse small and large intestinal explants. *Am. J. Physiol. Gastrointest. Liver Physiol.* **302**, G430–G438
  46. Schulz, B. L., Packer, N. H., and Karlsson, N. G. (2002) Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. *Anal. Chem.* **74**, 6088–6097
  47. Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362
  48. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906
  49. van der Post, S., Subramani, D. B., Backstrom, M., Johansson, M. E., Vester-Christensen, M. B., Mandel, U., Bennett, E. P., Clausen, H., Dahlen, G., Sroka, A., Potempa, J., and Hansson, G. C. (2013) Site-specific O-glycosylation on the MUC2 mucin protein inhibits cleavage by the *Porphyromonas gingivalis* secreted cysteine protease (RgpB). *J. Biol. Chem.* **288**, 14636–14646
  50. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860
  51. Johansson, M. E., Jakobsson, H. E., Holmen-Larsson, J., Schutte, A., Ermund, A., Rodriguez-Pineiro, A. M., Arike, L., Wising, C., Svensson, F., Backhed, F., and Hansson, G. C. (2015) Normalization of host intestinal mucus layers requires long-term microbial colonization. *Cell Host Microbe* **18**, 582–592
  52. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372
  53. Schilling, B., Rardin, M. J., MacLean, B. X., Zawadzka, A. M., Frewen, B. E., Cusack, M. P., Sorensen, D. J., Bereman, M. S., Jing, E., Wu, C. C., Verdin, E., Kahn, C. R., Maccoss, M. J., and Gibson, B. W. (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: Application to protein acetylation and phosphorylation. *Mol. Cell. Proteomics* **11**, 202–214
  54. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815