THE STRUCTURE OF HUMAN 4F2hc ECTODomain PROVIDES A MODEL FOR HOMODIMERIZATION AND ELECTROSTATIC INTERACTION WITH PLASMA MEMBRANE.

Joana Fort1,2*, Laura R de la Ballina1*, Hans E Burghardt1, Carles Ferrer-Costa3, Javier Turnay4, Cristina Ferrer-Orta2, Isabel Usón5, Antonio Zorzano1, Juan Fernández-Recio5,6, Modesto Orozco3, María Antonia Lizarbe4, Ignacio Fita2 and Manuel Palacín1

From 1Institute for Research in Biomedicine, Barcelona Science Park, Department of Biochemistry and Molecular Biology, Faculty of Biology, University of Barcelona, and CIBERER, Barcelona, Spain, 2Instituto de Biología Molecular de Barcelona (IBMB-CSIC), Institute for Research in Biomedicine, Barcelona Science Park, 3Molecular Modeling and Bioinformatics, Institute for Research in Biomedicine, Barcelona, Spain, 4Departamento de Bioquímica y Biología Molecular I, Facultad de Ciencias Químicas, Universidad Complutense, Madrid, Spain, 5ICREA at Instituto de Biología Molecular de Barcelona (IBMB-CSIC), Spain, 6Barcelona Supercomputing Center, Barcelona, Spain.

Running head: Structure of 4F2hc (CD98) ectodomain

Address correspondence to:
Manuel Palacín, Institut de Recerca Biomedica, Parc Científic de Barcelona, Josep Samitier 1-5, E-08028, Barcelona, Spain. Tel.: +34-934037199; Email: mpalacin@pcb.ub.es

and

Ignacio Fita, Institut de Recerca Biomedica, Parc Científic de Barcelona, Josep Samitier 1-5, E-08028, Barcelona, Spain. Tel.: +34-934034949; Email: ifrcri@ibmb.csic.es

Keywords: CD98, FRP-1, Heteromeric Amino acid Transporter, zinc coordination, heavy subunit, extracellular domain.

4F2hc (CD98hc, FRP-1, and SLC3A2) is a multifunctional type II membrane glycoprotein involved in amino acid transport and cell fusion, adhesion and transformation. The structure of the ectodomain of human 4F2hc has been solved using monoclinic (2DH2) and orthorhombic (2DH3) crystal forms at 2.1 and 2.8 Å, respectively. It is composed of a (βα)8 barrel and an antiparallel β8 sandwich related to bacterial α-glycosidases, although lacking key catalytic residues and consequently catalytic activity. 2DH3 is a dimer with Zn2+ coordination at the interface. Human 4F2hc expressed in several cell types resulted in cell surface and Cys109 disulfide bridge-linked homodimers with major architecture features of the crystal dimer, as demonstrated by cross-linking experiments. 4F2hc has no significant hydrophobic patches at the surface. The N-terminus of the solved structure, including position of Cys109 residue located four residues apart from the transmembrane domain, is adjacent to the positive face of the ectodomain. This location of the N-terminus and the Cys109-intervening disulfide bridge impose space restrictions sufficient to support a model for electrostatic interaction of 4F2hc ectodomain with membrane phospholipids. These results provide the first crystal structure of Heteromeric Amino acid Transporters and suggest a dynamic interaction of 4F2hc ectodomain with plasma membrane.

4F2hc (CD98hc, FRP-1 and SLC3A2) is a multifunctional type II membrane glycoprotein, involved in amino acid transport (1), cell fusion (2) and β1 integrin-dependent adhesion (3). 4F2hc and the homologous rBAT are the heavy subunits of the Heteromeric Amino acid Transporters...
(HATs), which are linked by a disulfide bridge to the catalytic light subunit (Fig 1A). One of six light subunits (LAT1, LAT2, y’LAT1, y’LAT2, asc-1 and xCT) heterodimerizes with 4F2hc, thereby rendering a range of transport activities. 4F2hc-associated light subunits are involved in human pathology (y+LAT1 mutations cause lysinuric protein intolerance and xCT is the receptor of Kaposi's sarcoma-associated herpesvirus (4-6)). The known role of the heavy subunits is to bring the holotransporter to the plasma membrane. Moreover, 4F2hc is involved in cellular transformation since it is highly expressed in tumor cells, its expression correlates with tumor development, progression and metastatic potential, and its over-expression leads to cell transformation (7-10). 4F2hc is a mediator of β1-integrin signaling (11). Recently, a metabolic activation-related CD147-4F2 complex has been identified on the cell surface that may play a critical role in energy metabolism, probably by coordinating the transport of lactate (via MCT1 and MCT4) and amino acids (via LAT1) (12). Integrin interaction and the CD147-4F2 complex may explain the role of 4F2hc in cellular transformation.

The role of the big ectodomain of the heavy subunits of HATs remains largely unknown. The 4F2hc ectodomain (4F2hc-ED) is required for plasma membrane localization of the light subunits LAT2 and y’LAT2 (13). 4F2hc-ED might also modulate β1 integrin function and tumorigenicity (10), though interactions with β1 integrins involve the transmembrane and the N-terminal intracellular segments of 4F2hc (14). Recently, interaction of galectin-3 with 4F2hc-ED has been reported to play a role in throphoblast formation by cell fusion in the placenta (15).

Intriguingly the ectodomains of the heavy subunits of HATs show homology with α-amylases (glycosyl hydrolase family 13 members) (16). The low amino acid sequence identity of the C-terminal half of 4F2hc-ED with α-amylases precluded the generation of a reliable 3D model by homology (1). The crystal structure of 4F2hc-ED solved here supports an electrostatic interaction of the ectodomain with the plasma membrane and suggests a receptor function for the α-amylase-like cleft. The crystal structure also evidences a tendency of 4F2hc to form homodimers, which has been confirmed in cell expression studies.

Materials and Methods

Production of human 4F2hc-ED in E. Coli. The cDNA sequence coding the 4F2hc-ED (residues 111-530 of human 4F2hc) was digested from the pSPORT-4F2hc plasmid (17) with NruI and HindIII. The insert was cloned into the pTrcHisA vector (Invitrogen), previously digested with BamHI, and then treated with Klenow and digested with HindIII. The resulting plasmid was transformed into BL21 (DE3) cells (Novagen) and the cells were grown in LB media at 37 °C. Cells with OD<sub>600</sub>=0.6 were induced with 1 mM IPTG followed by culture for 16 h at 37 °C. Harvested cells were resuspended in lysis buffer [20 mM potassium phosphate pH 8.0, 500 mM NaCl, 1 mg/ml lysozyme (Roche), 5 µg/ml DNAse (Roche), 1 mM PMSF, aprotinine, pepstatine and leupeptine] and lysed on a French Press at 20,000 psi. The resulting cell lysate was centrifuged twice at 16,000 g for 45 min. The soluble fraction was applied to a HiTrap Chelating 5 ml column (Amersham-Biosciences), previously charged with Ni<sup>2+</sup>, in an Äkta-FPLC. The buffer was changed to 50 mM Tris, 1mM CaCl<sub>2</sub>, 0.1% tween-20, and 4F2hc-ED was eluted by overnight digestion with enterokinase (50 U/column) (Invitrogen). Furthermore, it was purified by monoQ chromatography and the final yield was 2 mg of 4F2hc-ED/L initial culture. Protein was concentrated with centrifugal filters (Amicon Ultra 15 10,000 Da, Millipore) to ~15 mg/ml for crystallization. Circular dichroism was used to control the quality of the protein samples used in the crystallization trials. The percentage of secondary structure of 4F2hc-ED derived in solution was in agreement with the crystal structure (data not shown). A similar approach to produce human rBAT-ED was unsuccessful due to the low protein production and the incorporation into inclusion bodies (data not shown).

Crystallization and data collection. Two types of 4F2hc-ED crystals, needles and plates, were obtained by the hanging drop vapor diffusion method at 20°C. Needles were monoclinic (space group <i>P2</i><sub>1</sub>) and grew in 25% PEG 4000, 0.2 M sodium acetate and 0.1 M Tris buffer pH 8.5. Plates were orthorhombic (space group <i>P2</i><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>) and grew in 25% PEG 3350, 0.2 M ammonium sulfate, measured pH ~6. Crystals were flash-
cooled in the presence of 10% glycerol for the monoclinic crystal and 30% PEG 4000 1M KI for the orthorhombic crystal. X-ray diffraction data were collected at beamline ID13 at the ESFR (Grenoble) with a MarCCD detector at 100K. Data were processed and scaled using the DENZO and SCALEPACK programs, respectively (Table 1)

Structure determination and refinement. Molecular replacement searches on the monoclinic crystals, which contain one 4F2hc-ED molecule in their asymmetric unit, were performed following the CaspR protocols [CaspR server; CNRS, Marseille (18)] starting with three PDB structures of prokaryotic α-glycosidases (PDB IDs 1JI2, 1SMA and 1UOK), which had sequence identities of about 25% with respect to 4F2hc-ED. The TIM-barrel domain fragment from the best refined solution was then fixed and the C-terminal domain was reoriented using MOLREP (19). The new polyalanine structure was further modified by rigid body refinement with Refmac5 (20) of different fragments. The resulting model allowed us to obtain an initial solution, with AMoRe (21), for the orthorhombic crystal, which contains two 4F2hc-ED molecules in the asymmetric unit. Electron density was improved by averaging between the two crystal forms with DMMULTI (22). Model rebuilding and refinement was completed, for the two crystal forms, by iterative cycles of manual and automatic refinement with the graphic programs O (23) and Refmac5, respectively (Table 1). The stereochemistry of the models was checked with PROCHECK (24). The programs T-Coffee and ESPript were used for sequence and structural alignments and their representation, respectively (25, 26). Cavities were analyzed with the CASTp plug-in (27). Electrostatic surfaces were computed with GRASP (28). All structure pictures were drawn with PyMol (29).

Co-precipitation of 4F2hc-HA and His-4F2hc. HeLa cells were transiently co-transfected with combinations of wild-type human 4F2hc-HA, C109S 4F2hc-HA, pCDNA3.1 and wild-type human His-4F2hc (see Supplementary Materials and Methods). Two days later, cells were lysed with 400 µl/10 cm-plate of lysis solution [0.3 M NaCl, 1% Triton X-100, 0.025 M imidazole, protease inhibitor mixture (1 mM phenylmethylsulfonyl fluoride, 1 U/ml aprotinin, 1 µM leupeptin and 1 µM pepstatin), and 0.05 M NaPO₄; pH 7.4]. Cells were then scraped, transferred into a 1.5-ml tube, and incubated on a rotating orbital shaker at 4 °C. After 1 h, the insoluble material was removed by centrifugation at 12,000 rpm for 10 min. Supernatants were incubated with 20 µl of Ni-NTA-agarose beads (Qiagen) previously equilibrated with 150 µl of lysis solution for 30 min at 4 °C. After 1.5 h incubation at 4 °C with the solubilized cellular proteins, beads were washed twice in 750 µl of a PBS solution containing 0.5 M NaCl and 0.05 M imidazole (pH 7.4). Proteins were eluted by 15 min of shaking incubation at 25 °C with 45 µl of elution buffer (PBS containing 0.5 M NaCl and 0.25 M imidazole, pH 7.4). 4-fold concentrated Laemmli Sample Buffer (LSB) was added to the eluates from the Ni-NTA beads.

Cross-linking. Intact cells over-expressing different forms of 4F2hc-HA were grown to 90% confluence in 10-cm plates and then washed two times with PBS pH 7.0. Cross-linker BM[PEO]₂ (Pierce) (1 mM in PBS pH 7.0) was added for 1 h incubation at room temperature before termination with 50 mM dithiothreitol (DTT). After washing with PBS, cells were homogenized in lysis buffer (PBS, 1% Triton X-100 and the same protease inhibitor mixture as above) at 4 °C for 1 h. Lysates were centrifuged at 13,000 g for 5 min and supernatants were processed for Western blot analysis in presence of DTT (100 µM).

Endo H and endo F treatments. Total protein extracted as in the previous section (cross-linking) from HeLa cells (1 mg/ml) were treated with Glycoprotein Denaturing Buffer and incubated at 99 °C for 10 min. Reaction mixtures for Endo Hf (30 units/µl in G5 buffer) and PNGase F (30 units/µl in G7 buffer with 1% NP40) treatments were incubated at 37 °C for 1 h. Controls were performed with water instead of enzyme. In cases where reactions were carried out in non-reducing conditions, the Glycoprotein Denaturing Buffer was substituted by 0.5% SDS. All reagents and buffers were from New England Biolabs.

Western blot analysis. Protein extracts, obtained as above, from different origins in LSB containing or not 100 mM DTT were loaded for SDS–PAGE (7.5% polyacrylamide). c-Myc, HA, and His tags were detected with antibodies sc-789 (1:500, Santa Cruz Biotechnology), 3F10 (1:1000,
Roche) and anti-HisG (1:5000, Invitrogen), respectively.

Computational Analysis. For molecular dynamics simulations the monomeric crystal structure was fully solvated with an octahedron tip3p water box, ions were added and final system was minimized and equilibrated to 300K using amber force field (30) in namd program (31, 32). Finally, 10 ns of production were collected and analyzed as described (33). ODA calculations were performed by using a variation of the described protocol, defining the optimal patches from the center of coordinates of every residue side-chain instead of from a series of surface points as in the original method (34). Protein-protein docking calculations were performed by the pyDock protocol, as recently described (35). We used FTDOCK (36) to generate 10,000 rigid-body docking orientations, which were evaluated by the pyDock scoring function optimized for rigid-body docking. The resulting orientations were further clustered by in-house software. Briefly, for each docking solution, the rotation matrix and translation vector were converted to the rotation axis, rotation angle and translation vector that relate ligand and receptor coordinates. The combined 3 angular measures defined a distance between solutions that was used in an in-house implementation of a standard clustering method. Symmetry and filter distances [<17 Å between the two homologous S444C and S480C residues for cross-linking experiments and distance <30 Å between the two homologous residues Gln115 for disulfide bridge formation (residues previous to Gln115 were not included in the docking calculations due to their mobility)] were also applied using in-house software.

RESULTS

Overall structure of human 4F2hc-ED. We built a 4F2hc-ED construct that included the human 4F2hc residues from Glu111 to the C-terminus Ala529, plus a six-residue N-terminal extension (Lys-Asp-Arg-Trp-Gly-Ser) corresponding to the expression plasmid. The construct was crystallized in both a monoclinic (P2₁) and an orthorhombic (P2₁2₁2₁) crystal form. Structure determination was achieved by molecular replacement and crystal averaging (see Materials and Methods). The final refined structures present crystallographic agreement factors R and R_free of 17.8/23.0 and 22.0/27.5 at resolutions of 2.1 and 2.8 Å for the monoclinic and orthorhombic crystals, respectively (Table 1). Structures include residues Gly109-Ala529 for the subunit in the monoclinic crystal (PDB ID: 2DH2), and Ala114-Ala529 and Gly109-Ala529 for the two subunits in the orthorhombic crystal (PDB ID: 2DH3).

The structure of 4F2hc-ED can be described in terms of topology and multidomain organization of α-amylases (Figs. 1B and C). In 4F2hc-ED, the so-called A-domain in α-amylases, a (β/α)₈ TIM barrel, corresponds to the N-terminal globular domain. This domain starts with two contiguous tryptophane residues (Trp117-Trp118) without insertions until Lys431. A five-residue long helix (Arg433-His438) connects the A-domain to the carboxy-terminal domain, which corresponds to the so called C-domain, which, although presenting a variable β-sandwich topology, is always present in α-amylases. The C-domain of 4F2hc includes residues from Asp439 to the carboxy end Ala529. The interface between domains A and C is very hydrophobic, 67% of the total contact surface (2538 Å²), and involves mainly the secondary structure elements Aα₆ to Aα₈ and Cβ₁ to Cβ₃. 4F2hc lacks a B-domain, a globular insertion between the third β-strand (Aβ₃) and the third α-helix (Aα₃) of the A-domain, which is present in many α-amylases. In the three independent crystallographic 4F2hc-ED structures determined, residues that precede Trp117 adopt an extended conformation pointing towards the interface between the A- and C-domains in the vicinity of Trp456, a fully conserved residue in 4F2hc from vertebrates. Residue 109 in 4F2hc is a cysteine that makes a disulfide bridge with the light subunit and is just four residues away from the putative transmembrane segment. Our structure suggests that the location of the N-terminus imposes strong structural restrains with respect to both the docking of 4F2hc onto the membrane and the interactions with the light subunit in the heterodimer (Fig. 1B). Human 4F2hc presents four putative N-glycosylation sites in the A-domain (N264 in Aα₄, N280 just before Aα₈ and Cβ₁ to Cβ₃. 4F2hc lacks a B-domain, a globular insertion between the third β-strand (Aβ₃) and the third α-helix (Aα₃) of the A-domain, which is present in many α-amylases. In the three independent crystallographic 4F2hc-ED structures determined, residues that precede Trp117 adopt an extended conformation pointing towards the interface between the A- and C-domains in the vicinity of Trp456, a fully conserved residue in 4F2hc from vertebrates. Residue 109 in 4F2hc is a cysteine that makes a disulfide bridge with the light subunit and is just four residues away from the putative transmembrane segment. Our structure suggests that the location of the N-terminus imposes strong structural restrains with respect to both the docking of 4F2hc onto the membrane and the interactions with the light subunit in the heterodimer (Fig. 1B).
located near the 4F2hc-ED N-terminus giving free access of this face of the ectodomain to the plasma membrane surface.

The α-glycosidase-like cleft of human 4F2hc-ED. The catalytic A-domain between the structures of 4F2hc-ED (monoclinic crystal) and oligo 1,6-glucosidase from Bacillus cereus (PDB ID: 1UOK) gives a rmsd of 1.96 Å for 260 (83%) equivalent residues. Indeed, the structure of the catalytic A-domain is better preserved that what was anticipated on the basis of the low amino acid sequence identity, in particular for the C-terminal half, starting in Aα4, of the domain (Supplemetary Fig. 1). The active site of α-amylase family members is highly conserved with key catalytic residues located at the C-terminus of the β-strands from the A-domain: a nucleophile (Asp199 in 1UOK), a proton donor (Glu255) and a substrate-positioning residue (Asp329) (16). In 4F2hc, this site corresponds to a deep and wide cleft (Supplementary Fig. 2A), with a distinct shape to that of α-amylases and where only the residue acting as the nucleophile (Asp248) is conserved, although at a slightly different position (Supplemeytary Fig. 1 and 3B). Therefore, the structure indicates that apart from its general similarity to α-amylases, 4F2hc is not likely to have α-glycosidase activity. In fact, a detailed screening using D-glucose, D-galactose or D-mannose derivatives of 4-methylumbelliferone as substrates failed to detect α-glycosidase activity (Supplementary Fig. 2C).

Zinc binding and crystal homodimer. The two 4F2hc-ED subunits found in the asymmetric unit of the orthorhombic crystal are related by a local two-fold symmetry with the axis running almost parallel to the c crystal cell edge (Fig. 2A). Using X-ray fluorescence spectroscopy (data not shown), a high electron density peak at the interface of the two subunits forming the homodimer was found to correspond to a Zn\(^{2+}\) atom (Fig. 2A-C). This metal ion sits on the local two-fold axis and presents a tetrahedral coordination with three residues from the first subunit (Asp439, His441, both in Cβ1 and His455 in Cβ2) and only one (His441) from the second subunit (Fig. 2B-C). Therefore, coordination of Zn\(^{2+}\) breaks the homodimer symmetry precisely in the vicinity of the local two-fold axis. The unexpected binding of Zn\(^{2+}\), a trace metal not added during purification or crystallization procedures, suggests a high affinity (picomolar range) for this ion. In the 4F2hc-ED subunit determined in the monoclinic crystal, the side-chains of Asp439, His441 and His455 are exposed to the solvent, without participating in interactions with other protein subunits or in metal coordination (Fig. 2D). The homodimer presents a contact surface on each subunit of ~1200 Å with ten hydrogen bonds and a salt bridge between the two subunits.

Homodimerization of human 4F2hc in cultured cells. To check whether human 4F2hc homodimerizes in vivo, HeLa cells were co-transfected with two tagged versions of human 4F2hc (see supplementary Materials and Methods): 4F2hc-HA (HA tag at the C-terminus) and His-4F2hc (six His tag at the N-terminus) (Fig. 3A). Low molecular weight bands compatible with 4F2hc monomers (85-95 kDa) and high molecular weight bands (up to ~200 kDa) compatible with 4F2hc homodimers were detected with both tags in non-reducing conditions. After His-affinity purification, a HA-immunodetected high molecular weight band (170-185 kDa) was observed only in cells co-transfected with both 4F2hc-tagged versions. This indicates the formation of 4F2hc homodimers. These homodimers (D) were abolished in reducing conditions (Fig. 3B), where only monomers (M) and proteolytic fragments (P) were visible (Fig. 3B; and see Supplementary Fig. 4 legend). Human 4F2hc expressed in CHO-K1 cells or in Xenopus oocytes also showed DTT-sensitive bands with SDS-PAGE mobility compatible with the formation of homodimers (Supplementary Fig. 3). In contrast to HeLa cells (where we have not been able to visualize heterodimers) a band corresponding to the complex of 4F2hc and light subunits (~130 kDa) was detected in CHO-K1 cells expressing human 4F2hc and in oocytes co-expressing human 4F2hc and a tagged version of the light subunit xCT (myc-xCT) (Supplementary Fig. 3). These results demonstrate that over-expression of human 4F2hc in cultured cells results in the formation of DTT-sensitive 4F2hc homodimers.

Human 4F2hc contains two Cys residues. Cys330 has a lateral side-chain that is not accessible to the surface (located just before Aβ7
in the A-domain) and is not expected to be involved in any protein-protein disulfide bridge. Cys109 is the residue intervening in the disulfide bridge with 4F2hc-associated light chains (37). Mutation C109S (See Supplementary Materials and Methods) abolished the presence of 4F2hc homodimers in SDS-PAGE (Fig. 3B). In contrast, mutations that eliminate the Zn$^{2+}$ coordination patch identified in the crystal 4F2hc-ED homodimer (the triple mutant D439A, H441A and H455A; see Supplementary Materials and Methods) did not abolish homodimerization of 4F2hc in vivo (Fig. 3B). Similar results pointing to the formation of a homodimer of 4F2hc linked by a disulfide bridge through residue Cys109 and independent of Zn$^{2+}$ coordination were obtained in oocytes expressing human 4F2hc (Supplementary Fig. 3B). Interestingly, mutation C109S did not abolish 4F2hc homodimer formation. Thus, 4F2hc monomers were HA-immunodetected after His-affinity purification of solubilized material from cell expressing C109S 4F2hc-HA and His-4F2hc (Fig. 3A). This co-purification was specific since no 4F2hc-HA signal was detected in His-affinity purified extracts from cells expressing C109S 4F2hc-HA alone (Fig. 3A). This indicates the tendency to homodimerize of 4F2hc besides disulfide bridge formation.

Human 4F2hc homodimers are formed in the endoplasmic reticulum (ER) as demonstrated by its sensitivity to endo H treatment (Supplementary Fig. 4). Surface biotin labeling revealed expression of 4F2hc-HA monomers and homodimers at the plasma membrane, whereas no labeling was observed for an ER marker (BiP) (Supplementary Fig.5)

**Confirmation of the 4F2hc homodimer architecture in cultured cells.** The different contribution of Zn$^{2+}$ coordination to homodimerization of 4F2hc in the crystal and in vivo raises doubts as to whether the dimer found in the crystal corresponds to the homodimer structure in the cell. To verify this crucial point, we designed a series of cross-linking experiments (Fig. 3C) using BM[PEO]$_2$, a reagent that is membrane-impermeable and non-cleavable by reducing agents with the capacity to cross-link thiol groups 3.5 to 14.7 Å apart (38). BM[PEO]$_2$ cross-linked 4F2hc-HA, but not 4F2hc-HA C109S, into homodimers at the surface of HeLa cells (Fig. 3C). This suggests that, as expected (see above), Cys109 is a target for the cross-linker. Serines 444 and 480 are located in the C-domain between the secondary structures Cβ1 and Cβ2, and Cβ4 and Cβ5, respectively (Supplementary Fig. 1) and, on the basis of the crystal homodimer structure, are at a reasonable distance (around 14 Å) to be cross-linked by BM[PEO]$_2$ if mutated to Cys. Interestingly, the reagent cross-linked the double mutants 4F2hc-HA C109S S444C and 4F2hc-HA C109S S480C (see Supplementary Materials and Methods) into a complex with the SDS-PAGE mobility of 4F2hc-HA homodimers (Fig. 3C). These results strongly support the notion that the ectodomains of the 4F2hc homodimer in vivo and in the orthorhombic crystal share similar architecture.

**Modeling of 4F2hc-ED homodimers.** This modeling was performed to further discard the need of Zn$^{2+}$ coordination for 4F2hc-ED homodimerization. Optimal Desolvation Area (ODA) analysis (34) shows that the human 4F2hc-ED is almost absent of patches of low-desolvation energy at the surface (data not shown), which suggests that homodimerization is guided by shape and electrostatic complementarities rather than by the hydrophobic effect. Protein-protein docking simulations were performed in the absence of Zn$^{2+}$ coordination by taking two copies of the monomeric structure (PDB ID: 2DH2) of the protein as ligand and host. No restriction was made to enforce crystal-like packing. The docking solution ranked in position 2 by Pydock was very close to the pseudo-dimer X-ray structure (PDB ID: 2DH3) (Supplementary Table 1 and Supplementary Fig. 6), as well other top scored solutions. A final test was performed to check the possibility of alternative structures for the homodimer. For this purpose, the best 500 energy solutions were further filtered by the following criteria: 1) symmetry in the homodimer, and distance restrictions for 2) cross-linking experiments and 3) Cys109 disulfide bridge formation. These filters recover the information obtained from in vivo experiments but do not bias “per se” the structure towards the crystal. Application of these filters yielded 11 solutions, 7 of which clustered (Supplementary Table 1) in a
consensus structure close to the crystal solution (solution ranked 2 is shown in Fig. 4B,C). These calculations indicate that the dimer found in the crystal is a true conformation even in the absence of Zn\textsuperscript{2+} coordination. In all, the crystal structure, ODA analysis and docking simulations revealed electrostatic interactions as the driving force for homodimerization.

The surface charge distribution in monomeric and dimeric 4F2hc-ED. Electrostatic analysis showed that charged residues at the surface of 4F2hc-ED generate a macromolecular dipole: predominant negatively charged residues on top (face at the C-termini of the β-strands of the A-domain) and lateral sides, and predominant positively charge residues on the bottom side (face at the N-termini of the β-strands of the A-domain and where the N-terminus of the ectodomain is located) (Fig. 4A). This dipolar distribution is maintained in the homodimer of 4F2hc-ED, both in the orthorhombic crystal structure (PDB ID: 2DH3) and in the best solution of the cluster identified from virtual docking (Fig. 4B). The large positive patches and their well defined arrangement suggest that they might trigger the interaction of the protein with the plasma membrane.

DISCUSSION

This study provides the first crystal structure of a component of HATs. We previously attempted to model this ectodomain on the base of α-amylases homologous structures without succeeding (1). Once we have solved the structure of 4F2hc-ED, it does result compatible with other structures from α-amylase family. Despite the structural similarities found, the putative active site of 4F2hc presents major differences with respect to these enzymes. Accordingly, purified 4F2hc-ED showed no α-glycosidase activity for glucose-, galactose- and mannose-intervening substrates. In contrast, the derived structural model of the ectodomain of human rBAT suggests the presence of the main catalytic features found in α-amylases, including a B-domain, and consequently the conservation of glycosyl hydrolase activity (Supplementary Fig. 1).

The lack of this type of activity for 4F2hc-ED might indicate that this ectodomain still retains glycosidase-like binding properties. Other glycosidase-derived proteins such as the cytokines Ym1/2 (39, 40) and the ER degradation enhancing α-mannosidase I-like protein (EDEM) (41) have lectin activity. Unfortunately, a massive analysis performed using the Glycomics glycan chip (containing more than 260 glycans) failed to find any interacting glycan (data not shown; see Supplementary Materials and Methods). Possible binding of proteins or glycoproteins to the wide cleft of 4F2hc-ED is supported by the observation that some glycosidase inhibitors bind at the catalytic cleft and have Ig-like structures (e.g., Tendamistat (42)). Several proteins with this type of structure (e.g., galectin-3, ICAM-1, CD-147, CEA-CAM-1) have been proposed to interact with 4F2hc-ED (12, 15, 43-45).

In this study we have demonstrated the formation of 4F2hc homodimers linked by a disulfide bridge between Cys109 residues in several cell types expressing the human protein. Although 4F2hc homodimers have not been described previously, interpreting published data DTT-sensitive 4F2hc homodimer-compatible bands seem to be present in oocytes and COS cells expressing human 4F2hc and in human T lymphoblastoid cells (46-48). Homodimerization of 4F2hc could be facilitated by over-expression of the protein, but the tendency of the protein to dimerize is supported by the fact that the interconnecting disulfide bridge is not essential. The 4F2hc-ED homodimer found in the orthorhombic crystal most probably reflects the principal features of the homodimerization of 4F2hc in vivo, as only the putative N-glycosylation site N323 is near the homodimerization surface and the crystal homodimer is compatible with the restrictions imposed both by cross-linking experiments and by formation of the Cys109-intervening disulfide bridge. Most probably the interaction of the ectodomains in 4F2hc homodimers are of low affinity: this dimerization was observed only in one type of crystals and gel filtration of 4F2hc-ED revealed only the monomeric form (data not shown). This suggests that other parts of the protein (transmembrane domain and intracellular N-terminus) trigger homodimerization of 4F2hc.

Despite the unexpected presence of Zn\textsuperscript{2+} at the interface of the crystal homodimer, the coordination of this metal does not appear to be
required for 4F2hc homodimerization in vivo: i) the Zn$^{2+}$ coordination is not necessary for homodimerization of 4F2hc expressed in cells, ii) the residues involved are not conserved in mammals and iii) computational 4F2hc-4F2hc docking experiments yielded a cluster of solutions very similar to the crystal homodimer (without the Zn$^{2+}$ coordination), compatible with the binary symmetry, the Cys109 disulfide bridge formation and our cross-linking results. Nor does Zn$^{2+}$ coordination seem to play a role in the heterodimers. Mutation of the coordinating residues did not affect functional interaction of 4F2hc with light subunit xCT, even in the absence of the Cys109-intervening disulfide bridge (Supplementary Fig. 7). Thus, Zn$^{2+}$ coordination in human 4F2hc could be relevant for other protein interactions. Indeed, a very similar Zn$^{2+}$ coordination arrangement has been described in the superantigen Staphylococcus aureus enterotoxin A for high affinity interaction with MHC class II molecules (49).

Over-expression of 4F2hc in cells increases β1-integrin signaling (i.e., phosphatidylinositol 3-hydroxykinase, focal adhesion kinase and Akt) and this leads to cell transformation (50) and tumorigenesis in mice (10). The previously accepted view was that 4F2hc occurs in cells only as heterodimers (48). Here we show that over-expression of 4F2hc results in the presence of monomers and homodimers in the cell surface, which might have a role on the functions of 4F2hc in these conditions. In this sense, C109S mutation resulted in a maturation defect (i.e., lesser acquisition of complex N-glycosylation) of 4F2hc (Fig. 3), suggesting that homodimerization helps maturation and trafficking of 4F2hc. Thus, the Endo H-resistant form of C109S 4F2hc is underexpressed in HeLa cells (19 ± 4 % of total expressed 4F2hc versus 45 ± 2 % for wild-type 4F2hc; Student’s t test, p < 0.001 for n=6 independent transfections). Interestingly, C109S 4F2hc shows neither association with β1 integrin (51) nor cell transformation activity in cells over-expressing the protein (52). This correlation supports a role of 4F2hc homodimers in cell transformation.

Our results allow proposing a model for the disposition of 4F2hc homodimer in the membrane (Fig. 4B). No hydrophobic exposed patches were found in 4F2hc-ED surface. Thus, major hydrophobic interaction of 4F2hc with the membrane will be due to the highly hydrophobic transmembrane domain. The surface of 4F2hc-ED homodimer displays a marked polarized charge distribution with a positive surface in the N-terminus towards the Cys109 disulfide bridge and the transmembrane domain. This observation strongly suggests an electrostatic interaction of 4F2hc-ED with the phospholipid polar heads of the plasma membrane, characteristic of several peripheral and integral membrane proteins (53). This model might also provide clues about the arrangement of 4F2hc within the heterodimer in the plasma membrane since 4F2hc homo- and hetero-dimerization should be competitive processes because residue Cys109 participates in the disulfide bridge between monomers (this study) and in the heterodimers (37). In addition, the homodimer architecture suggests that the C-domain has a tendency for protein interaction. Then, external loops of the 4F2hc-associated light subunits might interact similarly with the C-domain, thereby favoring the docking of 4F2hc-ED onto the membrane. Moreover, the electrostatic nature of 4F2hc-ED interaction with plasma membrane might facilitate dynamic interaction with their multiple interacting proteins (e.g., light chains, β1 integrin complexes, CD147, galectin-3).

The structure of the 4F2hc ectodomain solved here seriously questions a catalytic role for this domain despite its overall similarities with α-amylases, brings the hypothesis of acting as a receptor for ligands, supports an electrostatic interaction with the plasma membrane and represents the first step towards the elucidation of the atomic structure of the heteromeric amino acid transporters. Moreover, the homodimerization of 4F2hc introduces a new element to understand the mechanisms involved in the functions of this protein in cell adhesion, cell transformation and tumorigenesis that need further investigation.

Acknowledgements

We thank Susanna Bial (Institute for Research in Biomedicine and CIBERER, Barcelona) for technical assistance. We also thank Chantal Abergel (CNRS, Marseille) who taught us how to use CaspR and Tania Yates for English editing of the manuscript. We acknowledge
Richard Alvarez and Angela Lee, from the Consortium for functional glycomics, for glycan binding screening. This study was supported in part by the Spanish Ministry of Science and Education (SAF2003-08940, BFU2006-14600 and BIO2005-06753), by the EC project grant 502802 EUGINDAT, and by La Marató-TV3. J.F. and L.R.B. are recipients of a predoctoral fellowship from the Spanish Ministry of Science and Education.

REFERENCES


**FOOTNOTES**

*J.F. and L.R.B. share first authorship*

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2DH2 and 2DH3).
FIGURE LEGENDS

Figure 1. Structure of 4F2hc-ED. A. HAT schematic representation. 4F2hc (pink) with a bulky N-glycosylated ectodomain (4F2hc-ED, covering residues W117 to the C-terminal A529) is linked by a conserved disulfide bridge (Cys109 in human 4F2hc) with a light subunit (blue), a 12 trans-membrane spanning non-glycosylated protein. Lateral (B) and upper (C) views of 4F2hc-ED structure. The N-terminal position corresponds to Cys 109. The structure is similar to that of α-glycosidases, including two domains: a TIM barrel (β/α)₈ and a C-terminal domain with eight antiparallel β-sheets.

Figure 2. Homodimer of 4F2hc ectodomain. A. Ribbons of the two molecules (chains A and B in green and yellow respectively) of 4F2hc-ED in the asymmetric unit of the orthorhombic crystal (PDB ID: 2DH3) with a zinc atom (purple sphere) in the interface zone. The homodimer symmetry axis is parallel to crystallographic axis c. B. Cartoon representation of the homodimer interface and tetrahedral coordination of zinc atom. C. Detail of 2.8 Å 2(Fo-Fc) density map (gray σ=1, pink σ=6) of zinc-coordinated residues. D. Stereo-view of 2.1 Å density map (σ=1) for the same residues in the monoclinic crystal where no density for the metal is present.

Figure 3. 4F2hc homodimerizes in HeLa cells. A. Empty pCDNA3.1 vector (φ), 4F2hc-HA (HA), 4F2hc-HA C109S (CS) and His-4F2hc (His) were transfected alone or in combination (HA+His; CS+His) in HeLa cells. Proteins were visualized with antibodies anti-His (αHis) and -HA (αHA) tags before (Input) and after Ni-NTA-agarose bead purification. B. Transient transfection with C-terminal HA-tagged 4F2hc (HA) in HeLa cells resulted in the formation of DTT-dependent duplex of bands with high molecular weight compatible with 4F2hc homodimers (D). In reducing conditions (+DTT), 4F2hc-HA appeared as duplex of bands of low molecular weight, compatible with the size of 4F2hc monomers (M) and proteolytic fragments (P). Empty pCDNA3.1 vector (φ) was used as a negative control. Mutation of Cys109 to Ser (4F2hc-HA C109S; CS) abolished the formation of disulfide bridge-linked homodimers, whereas disruption of the Zn²⁺ coordination patch by mutating all Asp439, His441 and His455 to Ala (4F2hc-HA D439A H441A H455A; Zn) had no effect on the formation of these homodimers. C. Cross-linking of human 4F2hc homodimers. Intact HeLa cells expressing different versions of human 4F2hc-HA (wild type, HA; C109S, CS; C109S S444C, 444; C109S S480C, 480) were treated with the sulphydryl cross-linker BM[PEO]₂ and processed for reducing Western blot. Anti-HA antibody revealed: before Endo F treatment (see Supplementary Materials and Methods), glycosylated homodimers (165-180 kDa), monomers (85-95 kDa) and proteolytic fragments (65-75 kDa) and after Endo F digestion, deglycosylated homodimers (D d, ~140 kDa), proteolized homodimers (D p, ~115 kDa), monomers (M d, ~70 kDa) and proteolized monomers (M p, ~55 kDa).

Figure 4. Electrostatic surfaces of 4F2hc-ED homodimer support a model for plasma membrane interaction. A. Electrostatic surface of monomer in the monoclinic structure (2DH2) has a dipolar distribution of charges with a positive patch in the N-terminal region (left) and negative surface residues in the other sites of the molecule (right). B. Electrostatic surface of N-terminal region of the homodimer from computational docking experiments. C. Model of 4F2hc homodimer interacting with membrane phospholipids. The positive face of the homodimer has a concave surface and 4F2hc-ED monomer is a rigid structure, as revealed by molecular dynamics analysis and circular dichroism (Supplementary Fig. 8). Therefore, two possibilities could be envisaged for the interaction of the ectodomain with the membrane. Either the ectodomain homodimer flattens to adapt to the membrane or deforms the phospholipid bilayer (shown in the figure), as it has been demonstrated for BAR domain homodimers (54).
Table 1. Data collection and refinement statistics (Molecular Replacement)

<table>
<thead>
<tr>
<th>PDB IDs</th>
<th>2DH2</th>
<th>2DH3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P₂₁,₂,₂₁</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>39.7, 67.8, 74.0</td>
<td>69.5, 101.8, 122.0</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.0, 98.5, 90.0</td>
<td>90.0, 90.0, 90.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>25.0-2.1 (2.2-2.1)*</td>
<td>30.0-2.8 (2.9-2.8)</td>
</tr>
<tr>
<td>R_sym</td>
<td>8.4 (54.9)</td>
<td>5.7 (27.0)</td>
</tr>
<tr>
<td>I/σI</td>
<td>7.3 (1.5)</td>
<td>15.7 (3.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.3 (98.8)</td>
<td>94.3 (96.2)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.5 (2.9)</td>
<td>7.1 (4.2)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>19.8-2.10</td>
<td>29.6-2.8</td>
</tr>
<tr>
<td>No. reflections</td>
<td>22,652</td>
<td>19,650</td>
</tr>
<tr>
<td>R_work / R_free</td>
<td>17.8/23.0</td>
<td>22.0/27.5</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3,269</td>
<td>6,504</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>4 (acetate)</td>
<td>1 (zinc)</td>
</tr>
<tr>
<td>Water</td>
<td>139</td>
<td>28</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>34.0</td>
<td>51.1</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>44.3</td>
<td>46.0</td>
</tr>
<tr>
<td>Water</td>
<td>37.5</td>
<td>36.8</td>
</tr>
<tr>
<td>R.m.s deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.012</td>
<td>0.006</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.446</td>
<td>1.180</td>
</tr>
<tr>
<td>Ramachandran Plot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residues in disallowed regions</td>
<td>1 (Ser 480)</td>
<td>1 in chain A (Phe218)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 in chain B (Leu 284)</td>
</tr>
</tbody>
</table>

*Values in parentheses are for the highest-resolution shell.
Figure 1

A

4F2hc-ED (W117-A529)

extracellular

Cys109

cytosol

B

C

by guest on September 1, 2017 http://www.jbc.org/ Downloaded from
Figure 2
Figure 3

A: Input and His-tagged purified samples. 

<table>
<thead>
<tr>
<th>kDa</th>
<th>αHis</th>
<th>αHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>121</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B: Samples with DTT and without DTT. 

<table>
<thead>
<tr>
<th>kDa</th>
<th>αHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>211</td>
<td></td>
</tr>
<tr>
<td>121</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

C: + DTT and + Endo F samples. 

- Endo F
+ Endo F

kDa
- HA
- CS
- 444
- 480

D: D d
P: P d
M: M d
D: D p
P: P p
M: M p
Figure 4
The structure of human 4F2hc ectodomain provides a model for homodimerization and electrostatic interaction with plasma membrane

Joana Fort, Laura R de la Ballina, Hans E Burghardt, Carles Ferrer-Costa, Javier Turnay, Cristina Ferrer-Orta, Isabel Usón, Antonio Zorzano, Juan Fernández-Recio, Modesto Orozco, María Antonia Lizarbe, Ignacio Fita and Manuel Palacín

J. Biol. Chem. published online August 26, 2007

Access the most updated version of this article at doi: 10.1074/jbc.M704524200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2007/08/27/M704524200.DC1

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2007/08/26/jbc.M704524200.citation.full.html#ref-list-1