

Published in final edited form as:

Nat Struct Mol Biol. 2016 February ; 23(2): 172–179. doi:10.1038/nsmb.3161.

Selective integrin endocytosis is driven by interactions between the integrin α -chain and AP2

Nicola De Franceschi¹, Antti Arjonen¹, Nadia Elkhatib^{2,3}, Konstantin Denessiouk⁴, Antoni G Wrobel⁵, Thomas A Wilson⁵, Jeroen Pouwels¹, Guillaume Montagnac^{2,3}, David J Owen⁵, and Johanna Ivaska^{1,6}

¹Turku Centre for Biotechnology, University of Turku, Turku, Finland ²Institut Gustave Roussy, Villejuif, France ³Institut National de la Santé et de la Recherche Médicale (Inserm), U1170, Villejuif, France ⁴Department of Biosciences, Åbo Akademi University, Turku, Finland ⁵Department of Clinical Biochemistry, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK ⁶Department of Biochemistry and Food Chemistry, University of Turku, Turku, Finland

Abstract

Integrins are heterodimeric cell-surface adhesion molecules comprising one of possible 18 α -chains and one of possible 8 β -chains. They control a range of cell functions in a matrix- and ligand-specific manner. Integrins can be internalised by clathrin-mediated endocytosis (CME) through β subunit-based motifs found in all integrin heterodimers. However, whether specific integrin heterodimers can be selectively endocytosed was unknown. Here, we found that a subset of α subunits contain an evolutionarily conserved and functional Yxx Φ motif directing integrins to selective internalisation by the most abundant endocytic clathrin adaptor, AP2. We determined the structure of the human integrin $\alpha 4$ -tail motif in complex with AP2 C- $\mu 2$ subunit and confirmed the interaction by isothermal titration calorimetry. Mutagenesis of the motif impaired selective heterodimer endocytosis and attenuated integrin-mediated cell migration. We propose that integrins evolved to enable selective integrin-receptor turnover in response to changing matrix conditions.

Introduction

Integrins are heterodimeric transmembrane receptors composed of an α - and a β -subunit. They mediate cell adhesion to the extracellular matrix (ECM) and link it to the cellular

Correspondence should be addressed to J.I. (Johanna.ivaska@utu.fi).

Author contributions

J.I. supervised the cell biology portion of the study, carried out experiments, analysed the data and wrote the manuscript with the contribution of N.D.F., G.M. and D.J.O. A.G.W., T.A.W and D.J.O. carried out the ITC and crystallography. N.D.F. conceived the study, designed, carried out and analysed most of the experiments with crucial help from A.A., N.E. and J.P. N.D.F. and K.D. carried out the evolutionary studies.

Competing financial interests

The authors declare no competing financial interests.

Accession codes: Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 5FPI.

cytoskeleton and signalling apparatus. Integrins regulate many physiological events, such as cell motility, cell survival, migration, proliferation and gene expression¹. In mammals, 24 different integrin heterodimers are expressed in a cell type- and tissue-specific manner, in which each heterodimer fulfils a specific biological role². The majority of matrix-binding integrins are composed of an α -subunit paired with the common $\beta 1$ -subunit, thus giving rise to a total of 12 integrin receptors. These bind to a repertoire of ligands and exhibit a certain degree of overlap in terms of binding specificity³. However, the signalling pathways activated by specific integrin heterodimers are distinct even when they are triggered by the same extracellular ligands. Thus, within subsets of integrins with overlapping ligand binding specificity, each receptor fulfils a distinct biological function. Integrin signalling occurs via the short, unstructured integrin cytoplasmic domains, and signalling specificity has been assigned to the α -subunit cytoplasmic domains. This is exemplified by the fact that swapping the α -cytoplasmic domains in integrins is sufficient to switch the intra-cellular signalling pathway triggered by the receptor, irrespective of the extracellular-matrix ligand engaged⁴ and integrin-mediated adhesion to collagen via distinct α -subunits triggers either attenuation of receptor-tyrosine kinase signalling or activation of a MAPK signalling cascade^{5,6}, on the basis of distinct protein-protein interactions at the α -cytoplasmic domain.

Integrins are constantly endocytosed and recycled back to the plasma membrane (PM) through multiple routes⁷. Tight regulation of integrin turnover from the cell surface is pivotal to a number of biological processes, including cell migration and cytokinesis^{8,9}, and has been implicated in cancer cell invasion and metastasis¹⁰. Integrins are predominantly endocytosed via clathrin-mediated endocytosis (CME) and this is governed by the intracellular domains of both α - and β -subunits. On the α -subunit, binding of the small GTPase Rab21 and p120RasGAP (RASA1) to the conserved membrane-proximal GFFKR-sequence, shared by all α -subunits, regulates endocytosis and recycling of the receptor, respectively^{11,12}. Conversely, monomeric clathrin adaptor components Dab2 and ARH associate with conserved segments shared by all β -integrin subunits and have been implicated in integrin endocytosis, whereas binding of the small GTPase Rab25 to β subunits regulates recycling in invasive protrusions^{9,13}. Thus, substantial advances have been made in our understanding of the mechanisms regulating integrin traffic. Nevertheless, the existence of regulatory pathways that would trigger preferential internalisation of one integrin heterodimer over another was unknown. Should they exist, such mechanisms would explain how cells rapidly and efficiently respond to the changing extracellular environment and orchestrate complex biological processes such as cell migration. Therefore, we aimed to identify a mechanism that allows cells to specifically regulate the traffic of a specific subset of integrins, which is one of the major outstanding issues in the field of integrin traffic.

The assembly polypeptide 2 (AP2) complex is a central player in CME and controls both the formation of clathrin-coated pits (CCP) and the recruitment of endocytic cargo. AP2 recruitment of receptors is mediated by the $\mu 2$ and $\sigma 2$ subunits of the complex. The $\mu 2$ subunit binds to "Yxx ϕ " motifs (in which x denotes any amino acid, and ϕ denotes a bulky hydrophobic residue) such as that found in the transferrin receptor¹⁴. Here, we identified a previously uncharacterized "Yxx ϕ " motif in a subset of integrin α -cytoplasmic domains and set out to investigate its evolutionary conservation, AP2 binding and functionality in integrin endocytosis on different matrices.

Results

A subset of integrin α chains carries a conserved Yxx ϕ motif

By aligning the membrane-proximal and cytoplasmic regions of all human integrin α -chains, we identified a putative Yxx ϕ internalization motif (in which ϕ can be leucine, isoleucine, methionine or valine) present in a subset of α -chains (Fig. 1a). We then separated integrins with overlapping ligand specificity into motif-containing and motif-lacking subgroups (Fig. 1b). The results suggested that the sequence might be important in differential integrin endocytosis on the same matrix. A related internalization motif, YxxG ϕ 15, is also present but it is embedded in the membrane and thus not accessible to cytoplasmic interactors. Because evolutionary conservation can define the functional importance of a specific sequence, we analyzed the evolutionary distribution of the integrin Yxx ϕ motif. The mechanism of endocytosis is evolutionarily ancient and the AP2 complex is present in almost all eukaryotes¹⁶, whereas integrins evolved later¹⁷. Our results showed that in addition to the membrane proximal GFFKR sequence, the most conserved region in all integrin α -chains, the Yxx ϕ motif, is highly conserved. The GFFKR motif in the α -integrin cytoplasmic domain was present in all metazoans, whereas the Yxx ϕ motif first emerged in *Branchiostoma floridae* (Cephalochordata) (Fig. 1c), thus suggesting a later appearance of this motif in evolution. Interestingly, the integrin I-domain, an important ligand-binding structural module present in a subset of integrin α -chains, appeared to have evolved only in Tunicata, later than the Yxx ϕ motif (Fig. 1c). These events predate the first diversification of the integrins from the more primitive receptor family found in Tunicata to the full set of human-like integrin α -subunits found in Clupeocephala, which belongs to a class of ray-finned fishes (Actinopterygii)¹⁷. Surprisingly, Clupeocephala shares similarities with humans in the number and types of integrin α -chains as well as in the distribution of the Yxx ϕ motif, with the exception of α 4 (ITA4) and α E (ITAE), in (Supplementary Fig. 1a). Laminin-binding integrins α 3, α 6 and α 7 all contain the motif (Fig. 1b). However, we unveiled a further element of complexity by aligning the splice variants of these integrins. These three laminin-binding integrins clustered together based on sequence similarity and ligand specificity¹⁸. Interestingly, they all had splicing isoforms showing modulation of the intracellular domain, exhibiting isoforms both with and without the Yxx ϕ motif (Supplementary Fig. 1b).

Sequence analysis on several mammals¹⁹ revealed that the cytoplasmic regions of some integrins (α M ITAM, α E ITAE, α X ITAX, α D ITAD) exhibited clear sequence variation, but the critical residues of the Yxx ϕ motif were strictly conserved, similarly to the GFFKR sequence (Supplementary Fig. 1c). In addition, even in lower organisms¹⁹ the Yxx ϕ motif exhibited stricter conservation compared to the C-terminal cytoplasmic domain in the α 2 (ITA2) and α 4 (ITA4) chains (Supplementary Fig. 1d). The conservation of the GFFKR motif parallels its functional importance in integrin activation²⁰ and integrin traffic⁸. A similar degree of conservation of the Yxx ϕ sequence suggests that this motif also has a critical role in integrin function.

Integrin $\alpha 2$ and $\alpha 4$ Yxx ϕ motifs interact with the AP2 complex

Prompted by the potential functional importance of the Yxx ϕ motif in the integrin α subunit we tested whether this sequence could mediate binding of integrin α -chains to the μ subunit of the AP complex. *In vitro* histidine-tagged AP2 $\mu 2$ C-terminal subdomain (C- $\mu 2$) (Supplementary Fig. 2a) exhibited clear binding to the $\alpha 2$ wild-type (WT) peptide (Fig. 1d). This interaction was significantly impaired upon mutation of the Y and ϕ residues ($\alpha 2$ -AxxA) (Fig. 1d, Supplementary Fig. 2b, c) or after incubation with a competing unlabeled $\alpha 2$ peptide (Supplementary Fig. 2b, c). (We were unable to test the full-length biotinylated $\alpha 4$ -integrin cytoplasmic peptide in this assay, owing to insolubility issues). Integrin peptides lacking the Yxx ϕ motif (including another collagen-binding integrin $\alpha 1$) demonstrated weaker binding to C- $\mu 2$ (Fig. 1d). We analyzed the strength of the α -integrin tail-C- $\mu 2$ interaction with Isothermal Titration Calorimetry (ITC), using an $\alpha 4$ peptide (QYKSILQE) and a longer $\alpha 2$ -integrin (WKLGFFKRKYEKMTKNPDEIDETTELSS). (Solubility issues precluded the use of a comparably short $\alpha 2$ peptide, as described below). These peptides showed clear, robust and reproducible binding to C- $\mu 2$ with a K_d of $\sim 70 \mu\text{M}$ for the $\alpha 4$ peptide (Fig. 1e) and $\sim 50 \mu\text{M}$ for the $\alpha 2$ peptide (Supplementary Fig. 2d), as compared with an $\alpha 5$ -integrin peptide (YKLGFFKRSLPYGTAMEKAQLKPPATSDA; Supplementary Fig. 2d), a $\beta 1$ -integrin peptide, and a control peptide of unrelated sequence, none of which showed any clearly detectable binding (i.e., $K_d > 300 \mu\text{M}$). By way of comparison, the mainly internally localized proteins with strong endocytic signals TGN38 (DYQRLN), transferrin receptor (SYTRFS) and epidermal growth factor receptor (FYRLMS) have K_d values for C- $\mu 2$ of $\sim 4 \mu\text{M}$, $\sim 15 \mu\text{M}$ and $\sim 20 \mu\text{M}$, respectively (primary data not shown). In all cases there was a large and similar enthalpic contribution to binding of between 35 and 40 kJ/mol, as would be expected because all peptides bind through key Y and ϕ residues and β -augmentation backbone hydrogen bonds, but differ in their entropic counterbalance, a factor which may result from different numbers of water molecules associated with the peptides and their flexibility. *In vivo*, all full length transmembrane proteins are embedded in a membrane, and there is hence less entropy to lose because the tails are essentially restricted to only lateral diffusion in the plane of the membrane. Therefore, although there may be a modest increase in affinity when the interaction is measured on a membrane rather than in solution, there should be a similar increase for all transmembrane cargo proteins. However, previous liposome-based surface plasmon resonance measurements of TGN38 immobilised in phosphatidylinositol 4,5-bisphosphate (PtdIns4,5P₂)-containing liposome membranes²¹ have shown a considerable increase in the apparent affinity of a Yxx ϕ signal for holo AP2, of approximately an order of magnitude, for the strongest known endocytic signal, TGN38, owing to the co-recognition of the cargo with multiple PtdIns4,5P₂ molecules (affinity for PtdIns4,5P₂ of $\sim 5 \mu\text{M}$).

In line with the *in vitro* data, in cells endogenous collagen receptor $\alpha 2$ integrin, but not $\alpha 1$ or $\alpha 5$ integrins, significantly associated with AP2 in proximity ligation assays (PLAs) with antibodies specific to the AP2 subunit α -adaptin and $\alpha 2$ integrin (Fig. 1f and Supplementary Fig. 2). Together, these data demonstrate that the Yxx ϕ motif in $\alpha 2$ and $\alpha 4$ integrins interacts with AP2 $\mu 2$ subunit.

Co-crystal of C- μ 2 and α 4-integrin Yxx Φ peptide

We solved the structure of C- μ 2 in complex with the Yxx Φ motif (QYKSILQE) from α 4 integrin to 2.8 Å resolution (data collection and refinement in Table 1). The details of the integrin-motif binding closely resembled those of previously characterized Yxx Φ endocytic motifs. C- μ 2 is composed of two interlinked β -sandwich subdomains (subdomain 1 strands 1-6 and 15-17, and subdomain 2 strands 7-14) (Fig. 2a). The peptide, for which there was good electron density for residues 2-7 (Fig. 2b,c), formed a β -augmentation with subdomain 1, packing next to strand (β 16) forming three backbone-to-backbone hydrogen bonds. The tyrosine and isoleucine side chains sat in complementary pockets primarily lined with side chains from residues on strands 1 and 16 (Fig. 2b,c), and all other visible peptide residue side chains pointed into the solvent. The walls of the tyrosine side chain pocket were created by Phe174, Lys203, Trp421 and Arg423, and the tyrosine of the motif formed a π -cation bond with Arg423 while the tyrosine's OH group formed a hydrogen bond with Asp176 at the base of the pocket. The presence of this negatively charged residue and the size of the pocket would be expected to preclude the binding of a phosphorylated tyrosine residue, owing to electrostatic repulsion and steric clashing, respectively (Supplementary Fig. 3a). This was indeed the case for the α 4-integrin motif – because the tyrosine-phosphorylated peptide (QY(p)KSILQE) (in which p indicates phosphorylation) showed no detectable binding to C- μ 2 in ITC (Fig. 1e) – as well as for the TGN38 motif in its tyrosine-phosphorylated form (D.J.O., unpublished data). The inability to bind a phosphorylated motif could provide a switch to regulate integrin endocytosis, although to date no published evidence exists for tyrosine phosphorylation at this site. The leucine side chain pocket was lined with aliphatic residues and, overall, the interaction of the α 4-integrin–C- μ 2 buried approximately 850 Å² of total solvent accessible-surface area, in accordance with its micromolar K_d .

In complexes between other proteins and juxtamembrane fragments of integrin α -chain tails, including the region of an α integrin corresponding to that in the complex with C- μ 2 described here, in integrin α -chain tails adopts either a helical^{22,23} or a largely unstructured conformation^{24–26}. According to our structure, to bind C- μ 2 the α -integrin tail must be in an extended conformation and must be accessible from the cytosol. This is in line with a study showing that a serine-phosphorylated α 4-tail, which does contain a Yxx Φ motif, adopts an extended conformation when in complex with 14-3-3 ζ ²⁶. The serine phosphorylated α 4-integrin tail in this structure and the C- μ 2-bound unphosphorylated α 4 peptide were conformationally very similar (Fig. 2d), even though their respective binding partners exhibited entirely different folds. Given that the intervening x residues of the Yxx Φ motif pointed away from the C- μ 2 surface into solvent, phosphorylation of the Y + 2 serine (Ser1011) in the α 4-integrin sequence QYKSIL would probably have little effect on the affinity of the motif for AP2 (Fig. 2e). In conjunction with our observation that tyrosine phosphorylation of the Yxx Φ motif in the α 4 tail inhibited the interaction with C- μ 2 (Fig. 1e), these data suggest that phosphorylation events at different sites could potentially function as integrin switches regulating endocytosis and signalling in cells. However, this will need extensive further investigations in the context of the full-length integrin in cells.

Yxx ϕ motif controls integrin recruitment to CCPs

CME is a major pathway for integrin endocytosis^{8,27} and cargo recruitment to CCPs on the PM can be visualized in live cells with expression of fluorescently tagged clathrin light chain (CLC) or AP2 μ ²⁸. Through total internal reflection fluorescence (TIRF) microscopy, we found that endogenous α 2 was transiently recruited to AP2 μ -mCherry puncta on the ventral sides of cells (Supplementary Fig. 3b and Supplementary Movie 1). To investigate this further, we co-expressed wild-type or Yxx ϕ motif-mutated (Supplementary Fig. 3c) GFP-tagged α 2 and α 4 integrins with AP2 μ -mCherry in cells. We selected these two integrins for investigation because they recognize different ligands and have biologically distinct functions and expression patterns. WT and AxxA mutant were expressed at similar levels on the cell surface for both α 2 and α 4 constructs (Supplementary Fig. 3d, e). Interestingly, WT α 2 and α 4 integrins, but not their mutant counterparts, colocalized with AP2 μ -mCherry (Fig. 3a). α 4 WT also localized to focal adhesions, whereas such a localization was limited for α 2 WT (Fig. 3a). However, the α 2 AxxA mutant, compared with WT, exhibited increased localization to focal adhesions (Fig. 3b), thus suggesting that reduced integrin endocytosis could indirectly result in increased integrin localization to focal adhesions. α 2 was also markedly endosomal compared to the AxxA mutant (Fig. 3c) and overlapped partially with Rab5, Rab7 and most prominently with Rab21 vesicles, in line with the important role of Rab21 in integrin endocytosis²⁹. Overlap with Rab11 and the endoplasmic reticulum were minimal (Supplementary Fig. 4). α 2-GFP, particularly the AxxA mutant, also localized to intracellular tubular structures, which partially co-localized with a Rab4-positive compartment (Supplementary Fig. 4). Integrins are endocytosed to vesicles that move along microtubules at an average velocity of 0.5-2 μ m/sec²⁹. α 2 WT positive vesicles showed similar dynamics, with abrupt accelerations followed by long pauses and an average speed of 0.9-4.6 μ m/sec. In contrast, α 2 AxxA was predominantly localized to the PM (Fig. 3c and Supplementary Movies 2 and 3). In cells spreading on collagen, α 2 WT showed dynamic vesicular-tubular movements supportive of cell spreading (Supplementary Movie 4), whereas α 2 AxxA was present mainly in the cell periphery on the PM (Supplementary Movie 5). Together, these data show that mutagenesis of the Yxx ϕ motif on α 2 and α 4 integrins alters receptor dynamics on the PM.

Yxx ϕ motif mutation impairs α 2 and α 4 integrin endocytosis

We found that AP2-mediated CME was important for α 2-integrin endocytosis. The presence of specific short interfering RNA (siRNA) targeting of AP230 resulted in the downregulation of the AP2 complex as a whole, as shown by α -adaptin staining (Supplementary Fig. 5a,b) and inhibited endocytosis of endogenous α 2 integrin significantly (Fig. 4a; transferrin receptor was used as a positive control⁷). Concordantly, integrin α 2 or α 4 AxxA mutants displayed significantly impaired endocytosis when compared with the WT receptor, as detected with a well-established biotin-based endocytosis assay⁷ (Fig. 4b,c). The conserved Y in the motif was the most critical residue, because mutating it alone was sufficient to disrupt the motif and to inhibit α 2 endocytosis (α 2 AxxM) to a similar extent than the AxxA mutation (Supplementary Fig. 5c). These effects were predominantly due to impaired endocytosis rather than altered integrin recycling, because inhibition of recycling with primaquine³¹ did not abolish the difference between internalization of α 4 AxxA-mutant receptor and internalized α 4 WT (Supplementary Fig. 5d). The inability of AP2 silencing or

disruption of the motif to fully abolish integrin endocytosis is in line with the ability of $\alpha 2$ integrin to be endocytosed via different pathways including AP-2 independent CME11 and Numb-dependent and Dab2-dependent pathway from the dorsal side of cells^{32,33}. For $\alpha 4$ integrin the internalization routes remain poorly investigated. Integrin activity influences the trafficking kinetics of $\beta 1$ -integrins⁷. In line with this finding, collagen engagement of $\alpha 2$ integrin increased endocytosis of $\alpha 2$ -GFP by $40\% \pm 28\%$ when compared to $\alpha 2$ -GFP endocytosis on fibronectin (Fig. 4d). Interestingly, the difference between endocytosed WT and AxxA-mutant $\alpha 2$ -integrin levels remained unaltered by collagen adhesion, thus suggesting that the Yxx ϕ motif is critical especially for the uptake of unengaged integrins in cells (Fig. 4d). Accordingly, silencing of AP2 reduced the endocytosis of inactive $\alpha 2\beta 1$ integrin heterodimer by 60% without a significant effect on active $\alpha 2\beta 1$ endocytosis (Supplementary Fig. 5e). According to our data, the $\alpha 5$ tail does not interact with C- $\mu 2$, although it contains a Yxx ϕ motif-like sequence (YxxAM). Accordingly, we found that mutagenesis of $\alpha 5$ -integrin (AxxxA mutant) had no effect on receptor endocytosis (Supplementary Fig. 5f). Together, our results indicate that C- $\mu 2$ -mediated integrin endocytosis is highly selective for Yxx ϕ -motif-containing integrins and predominantly regulates endocytosis of inactive integrins.

Introduction of Yxx ϕ to αV integrin stimulates endocytosis

To further investigate the role of the C- $\mu 2$ binding Yxx ϕ motif in integrin endocytosis, we introduced the WT (YEKM) or mutant motif (AxxA) from $\alpha 2$ integrin into αV integrin, which lacks the Yxx ϕ motif, either by replacing four or two of the endogenous residues with the motif sequences (Fig. 5a). Strikingly, introduction of the WT motif induced αV -CFP endocytosis significantly compared to the nonmutated αV -CFP, and insertion of the AxxA motif fully abolished αV -CFP uptake (Fig. 5a). Interestingly, the gain-of-function (GOF) effect was even more prominent in the αV chimera in which two instead of four αV -residues were replaced (Fig. 5a). Given that we did not test the affinity status of these GOF-integrins, we cannot rule out the possibility that their increased endocytosis could be linked to altered activity. Together, these data demonstrate that the Yxx ϕ motif in integrins is functionally important for AP2 binding and integrin endocytosis.

AP2 regulates integrin endocytosis via the Yxx ϕ motif

As the AP2 complex regulates CME both by direct cargo recruitment and by mediating the formation of CCPs as a whole¹⁴, we investigated whether decreased integrin endocytosis was the result of a loss of the AP2 μ -integrin Yxx ϕ motif interaction or of a general effect of AP2 depletion on CME. We generated stable cells lines expressing siRNA-resistant myc-tagged AP2 μ WT or AP2 μ F174A D176S mutant (AP2 Mut which is unable to bind the Yxx ϕ motif but is fully functional in other aspects of CME³⁰) and assessed integrin endocytosis in these cells. As previously reported³⁰, endogenous AP2 (lower band) was downregulated upon stable expression of exogenous AP2 μ subunits (Supplementary Fig. 5g) and was completely absent after siRNA silencing (Supplementary Fig. 5h). Importantly, expression of the AP2 μ F174A D176S mutant failed to support $\alpha 2$ WT and $\alpha 4$ WT endocytosis, whereas cells expressing AP2 μ -myc WT effectively endocytosed the WT integrins (Fig. 5b,c). As expected, the endocytosis of AxxA-mutant integrins was not supported by the expression of either AP2 μ WT or the AP2 μ F174A D176S mutant (Fig.

5b,c). These data demonstrate that specific interaction between AP2 and the integrin Yxx ϕ motif is functionally important for the CME of α 2 and α 4 integrins.

Yxx ϕ mutagenesis impairs α 2 and α 4 integrin functions

Integrin traffic has been implicated in cell spreading^{11,34} and is fundamentally important for efficient cell migration²⁷. To assess the effect of the Yxx ϕ -motif mutation on cell migration, we expressed α 2- or α 4-GFP WT or their AxxA mutants in cells lacking both endogenous collagen receptors (including α 2 integrin) and α 4 integrin (GD25 β 1A cells)^{35,36}, and monitored cell migration on collagen or the α 4-integrin specific ligand H120 (fibronectin fragment). Owing to the presence of plasma fibronectin in serum (which would complicate the collagen or H120 dependence of the assay by providing a ligand for endogenous α 5 β 1 and α v β 3 integrins expressed in these cells³⁶), we performed these assays in the absence of serum and in the presence of 10 ng/ml HGF, which supports integrin trafficking and stimulates cell migration³⁷. Mutation of the Yxx ϕ motif in α 2 and α 4 integrins strongly influenced cell shape and motility, such that cells expressing the mutant integrins adhered with multiple prominent adhesive protrusions (Fig. 6a) and were severely impaired in cell migration on collagen and on H120 (Fig. 6b). Importantly, these differences in migration were not due to different cell-surface expression levels (Supplementary Fig. 5i,j). In more migratory cells, we also observed the presence of retraction fibers when adhesion was mediated via α 2 AxxA mutant on Collagen I (Supplementary Fig. 5k). These data are in line with a reduced ability of AxxA mutant integrins to be endocytosed. These data suggest that integrin endocytosis is linked to cell motility, potentially via indirect effects on adhesion-site disassembly, as previously demonstrated^{9,38}, and that the ability of AP2 to recruit a subset of integrins as cargo is critical for the normal function of these integrins on their specific matrix ligands.

Discussion

Here we identified a previously uncharacterized conserved Yxx ϕ motif present in a subset of integrin α subunits and demonstrated its functional importance in α 2- and α 4-integrin endocytosis. We propose that the Yxx ϕ motif in the integrin superfamily is a crucial and ancient molecular code aimed at modulating endocytosis and possibly other trafficking events within the subcellular trafficking machinery. Indeed, the AP superfamily of adaptors includes at least five members mediating trafficking from different subcellular membrane compartments, and they all recognize the same motif sequence³⁹. Thus, this motif may function as a regulator of integrin traffic at multiple distinct steps. Therefore, the discovery of this motif on integrin can potentially open a new perspective into selective regulation of integrin traffic.

In conjunction with the evolutionary diversification of integrins, our findings provide insight into how a mechanism occurred to fine-tune their endocytosis to be specific to only one heterodimer out of several that are expressed in the same cell types at the same time and bind to the same matrix ligand. The ability of a cell to use the Yxx ϕ motif to discriminate between different integrin heterodimers at the level of endocytosis is likely to be fundamentally important, because different α / β 1-integrin heterodimers with overlapping

ligand specificity have fundamentally different signaling and cytoskeleton-regulatory functions in cells^{5,6}.

Tetraspanin CD151 is an integrin-binding protein with a Yxx Φ motif. Mutation of this motif, however, has no effect on the overall integrin endocytosis rate, although it decreases the rate of endocytosis of the CD151-associated integrin⁴⁰. Several proteins binding to the conserved GFFKR motif in the integrin α subunit have been characterized in the last few years^{20,41}. Two such proteins are Rab21 and RASA1, which bind the GFFKR motif in a mutually exclusive fashion and mediate integrin endocytosis and recycling, respectively¹². These proteins would be in proximity to the Yxx Φ motif when in complex with the integrin α -chain, and they may have synergistic or antagonistic effects on integrin function in adhesion and traffic. Integrins undergo complex cross-talk with multiple kinases^{42,43}. We found here that tyrosine phosphorylation of the Yxx Φ motif is incompatible in vitro with C- μ 2 binding. This implies the possibility of an additional level of regulation of integrin endocytosis in cells, whereby tyrosine kinases capable of phosphorylating the motif-containing integrin α -tails could regulate an endocytosis switch and potentially induce retention of a phosphorylated subset of integrins at the plasma membrane. Given that focal adhesions are hubs for active tyrosine kinases, adhesion signaling might regulate endocytosis via such a phosphorylation-dependent switch. This, however, remains to be investigated in cells.

Several trafficking regulators binding to the integrin β -tail have been described⁴¹. Numb and Dab2 interact with the membrane-proximal NPXY motif on the integrin β subunit⁴⁴. Interestingly, both Numb and Dab2 interact with AP2^{45,46}. Therefore, describing the nature of these interactions in relation to AP2 binding will be a stimulating avenue to explore and will be critically important to the understanding of how integrin heterodimer-specific endocytosis occurs and how it is regulated within the numerous overlapping protein-protein interactions in the adhesome.

Methods

Evolutionary analysis

Through the Conserved domain Database⁴⁷, the sequence of the transmembrane and cytoplasmic region of integrin α was found to be the Pfam⁴⁸ superfamily member Integrin_alpha pfam00357 (InterPro entry: IPR018184). Initial analysis of the integrin- α cytoplasmic regions was made with the NCBI conserved-domain search, which identified 407 sequences from 77 species with the Integrin_alpha pfam00357 signature sequence, and 12 known structures. The taxonomy of these sequences was analyzed manually. For the complete analysis, we extracted all (3,856) sequences from the UniProt database that belonged to the integrin α -chain family, among which there were 1,171 reviewed sequences, according to UniProtKB/Swiss-Prot (the manually annotated and reviewed section of the UniProt Knowledgebase (UniProtKB)), and 2,685 unreviewed sequences, according to UniProtKB/TrEMBL. All incomplete sequences showing lack of the transmembrane cytoplasmic regions were then discarded. All the remaining sequences were analyzed, and the presence of the Yxx Φ motif was identified in every sequence by analyzing the segment immediately adjacent to the GFFKR transmembrane motif. The existence of the putative I

domain was identified by the subsequent analysis of every UniProtKB integrin α -chain sequence through NCBI Conserved Protein Domain analysis. Sequences potentially containing the I domains in their structure did show Von Willebrand factor type A (vWA; cd00198) domain consensus signatures, whereas those that did not showed only the existence of separate FG-GAP (β -propeller; smart00191) repeats. The taxonomy of all identified sequences was analyzed both manually and with the UniProtKB database taxonomy tools. Additional missing annotation was performed with the Ensemble automatic gene annotation system⁴⁹. In Figure 1a, UniProt entry codes are P17301, P13612, Q13797, P11215, P20702, Q13349, P38570, P23229, Q13683, P26006, P56199, P08648, P20701, P06756, P53708, P08514, O75578, Q9UKX5.

Antibodies and labeled transferrin

Antibodies specific to the following proteins were used in this study. Primary antibodies: GFP (A11122, Molecular Probes), β -tubulin (ab6160, Abcam), integrin $\alpha 2$ (MCA2025, AbD Serotech), integrin $\alpha 2$ (EPR5788, Abcam), integrin $\alpha 2$ (AB1936, Millipore), active integrin $\beta 1$ (12G10; Ab30394, Abcam), inactive integrin $\beta 1$ (4B4; 6603113, Beckman Coulter), AP2 μ (EP2695Y, Novus Biological), and AP2 α -adaptin (MA1-064, Pierce) and transferrin AlexaFluor568-conjugated (T-23365, Invitrogen). Secondary antibodies: AlexaFluor488 and AlexaFluor568-conjugated IgGs (A21202, A21206, A10042, A10037 (immunofluorescence), Invitrogen), horseradish peroxidase (HRP)-conjugated anti-biotin antibody (7075, Cell Signaling Technology), and HRP-conjugated IgGs (NA931, NA934 (immunoblotting), GE Healthcare). Validation for all antibodies can be found on the manufacturers' websites. In addition, the antibodies specific to AP2 adaptin and AP2 μ were validated in this manuscript by silencing as shown in the figures. Original images of immunoblots used in this study can be found in Supplementary Data Set 1.

Plasmids

The following plasmids were used in this study: pMWH₆C- $\mu 2$ (160-435)50, Myc- $\mu 230$, Integrin $\alpha 2$ -GFP29, Integrin $\alpha 4$ -GFP51, integrin $\alpha 5$ -GFP52, dsRed-Rab753, RFP-Rab553, dsRed-Rab2129, pPS-CFP2- αV -Integrin (Addgene, 57212), mCherry-Rab4a (Addgene, 55125), mCherry-Rab11 (Addgene, 55124) Point mutations were done by QuikChange II Mutagenesis kit (Agilent technologies) according to manufacturer's instructions. All constructs were checked for integrity and correctness with sequencing.

TIRF microscopy

For total internal reflection fluorescent microscopy (TIRF-M), HeLa cells seeded onto a collagen- (for $\alpha 2$ -GFP expressing cells) or fibronectin- (for $\alpha 4$ -GFP expressing cells) coated glass-bottom dish were transfected with the indicated constructs and imaged the next day through a 100x 1.49 NA TIRF objective on a Nikon TE2000 (Nikon France SAS, Champigny sur Marne, France) inverted microscope equipped with a QuantEM EMCCD camera (Roper Scientific SAS, Evry, France / Photometrics, AZ, USA), a dual output laser launch which included 491 and 561 nm 50 mW DPSS lasers (Roper Scientific), and driven by Metamorph 7 software (MDS Analytical Technologies). A DV2 beam-splitter system (Roper Scientific / Photometrics) mounted on the light path enabled the simultaneous acquisition of the two emission channels. A motorized device driven by Metamorph allowed

the accurate positioning of the illumination light for evanescent wave excitation. Mann-Whitney test was used for statistical analysis.

Cells and Transfections

HEK293 cells were grown in DMEM supplemented with 1% penicillin-streptomycin, 10% FBS, and 1% L-glutamine. HeLa cells were grown in DMEM supplemented with 1% penicillin-streptomycin, 10% FBS, 1% Na Pyruvate, 2% HEPES, 1% non-essential amino acids and 1% L-glutamine. Plasmid and siRNA transfections were performed with Lipofectamine 2000 (Life Technologies) and Hiperfect (Qiagen), respectively. HEK293 cells stably expressing AP2 μ rescue constructs were obtained by selection with 250 μ g/ml of G418 (Sigma Aldrich). The MDA-231 stable cell line was grown in DMEM supplemented with 10% FBS and 0.5% geneticin. The GD25- β 1A cells are described in ref. 36. They were grown in DMEM Glutamax supplemented with 10% FCS and 10 μ g/ml of puromycin. Unless indicated otherwise, cell lines are from American Type Culture Collection (ATCC). All cells were routinely tested for mycoplasma and found to be mycoplasma free.

FACS analysis

FACS staining was performed as described earlier⁵⁴. Briefly, transfected HEK293 or HeLa cells were detached, fixed, washed with PBS and stained with primary antibodies against α 2 integrin (MCA2025, Serotec, 1:100) or α 4 (MAB16983, Millipore, 1:100) or with secondary antibody only in control cells for 1 h. Cells were then washed with PBS and stained with AlexaFluor647-conjugated secondary antibody (ab150107, Invitrogen, 1:400). After being washed, cells were suspended in PBS, and fluorescence was analyzed with flow cytometry (FACSCalibur, BD). Cell-surface integrin levels (AlexaFluor647 signal) were analyzed from the GFP-integrin-positive cells.

Structure determination

Residues 160-435 of rat μ 2 (C- μ 2) purified as described previously⁵⁰ and cocrystallized with the QYKSILQE peptide from integrin α 4 by hanging-drop vapor diffusion against 2.2 M NaCl, 0.4 M NaK phosphate, 20% (v/v) glycerol, 0.1 M MES, pH 6.5, and 5 mM DTT. Crystals were of space group *P*6₂ (unit-cell dimensions 125.6 Å, 125.6 Å and 74.4 Å) and diffracted to 2.8-Å resolution. Data were collected at 100 K on a RIGAKU rotating anode and were processed and scaled as described in ref. 51. The structure was solved by molecular replacement with PHASER and refined to final R and R_{free} of 18.2% and 23.2% with sequential rounds of refinement in REFMAC and manual rebuilding.

Isothermal titration calorimetry

C- μ 2 was gel filtered into 100 mM HEPES, pH 7.5, 500 mM NaCl and 1 mM TCEP. Peptides were dissolved in the same buffer. Experiments were performed with a Nano ITC from TA Instruments. C- μ 2 at 100 or 300 μ M was placed in the cell at 10 °C, and peptides at concentrations from 1 to 3 mM, depending on peptide, were titrated in with 24 injections of 2 μ l with injections separated by 5 min. A buffer into protein blank was subtracted from all data, and a minimum of at least four independent runs that showed clear saturation of binding were used to calculate the mean K_d of the binding reaction, its stoichiometry and

the corresponding s.e.m. values. Analysis of results and final figures were carried out with the NanoAnalyze Software.

Synthetic Peptides and Recombinant Proteins

Synthetic biotinylated peptides corresponding to the cytoplasmic domains of human ITGA1, ITGA2 WT, ITGA2 AxxA mutant, ITGA4, ITGA5 and ITGB1 or the unlabeled competing ITGA2 WT peptide were purchased from LifeProtein. Recombinant histidine-tagged AP2 μ C-terminal subunit (residues 158 to 435) used in pulldown assays was produced in *Escherichia coli* strain Rosetta BL21(DE3) and purified according to the manufacturer's instructions (BD Biosciences).

For crystallization and ITC, the following unlabeled peptides were used: TGN38 (DYQRLN), integrin α 4 (QYKSILQE), phosphorylated integrin α 4 (QpYKSILQE), α 2 integrin (WKLGFFFKRKYEKMTKNPDEIDETTELSS), α 5-integrin peptide (YKLGFFKRSLPYGTAMEKAQLKPPATSDA), transferrin receptor (SYTRFS), EGFR (FYRLMS) and unrelated control peptide (RMSQKIRLLSE).

Pull-down assays

For pulldown assays, equimolar amounts of biotinylated peptides were incubated with equal amounts of recombinant C- μ 2 in TBSt buffer plus 3% BSA at 4 °C for 2 h. Streptavidin-Sepharose beads were added and incubated 1 h at 4 °C. After the beads were washed, the bound protein was detected by western blot analysis. In competition experiments, a ten-fold molar excess of unlabeled peptide compared to the biotinylated peptide was added.

Immunofluorescence-based and PLA endocytosis assay

Silenced cells were plated on μ -Slide 8 well (Ibidi) coated with collagen I (5 μ g/ml). Slides were cooled on ice and cell-surface α 2 integrins or Transferrin receptor were labeled either with indicated anti- α 2 and secondary AlexaFluor488-conjugated antibodies or AlexaFluor561-labeled transferrin, respectively, in serum-free cell culture medium. Staining medium was aspirated, washed and replaced with fresh culture medium at 37°C. After internalization, plates were lifted on ice. The AlexaFluor488 fluorescence on cell surface was quenched by adding anti-AlexaFluor488 antibody as described in ref. 7 and incubation for 1 h on ice or with an acid-wash protocol which was performed as described in 55. Cells were fixed in 4% paraformaldehyde (PFAH) for 20 min at room temperature, and samples were imaged by confocal microscopy. For the PLA endocytosis assays, control or AP2-silenced cells growing on 5 μ g/ml collagen were cooled on ice and labeled with integrin α 2 and inactive (4B4) or active (12G10) integrin β 1 in serum-free medium. After washing, endocytosis was permitted for 15 min at 37 °C in serum-free medium, after which the cells were transferred back to ice. Subsequently, cells were fixed and permeabilized, and the PLA reaction was carried out according to the manufacturer's standard protocol (Duolink, Sigma-Aldrich) with specific secondary antibodies recognizing the anti- α 2 antibody or the anti- β 1 integrin antibody.

Biotin-based endocytosis assay

Biotin-based endocytosis assays were performed as described previously⁷. In brief, cells transfected with the fluorescent protein-tagged integrins (WT of mutant) were placed on ice and washed once with ice-cold PBS, and surface proteins were then labeled with cleavable NHS-SS-biotin (Pierce) in PBS at 4 °C for 30 min. Labeled cells were subsequently transferred to prewarmed serum-free medium, and internalization was allowed at 37 °C for the indicated times. Remaining surface biotin was removed by MesNa cleavage at 4 °C for 30 min; this was followed by quenching with iodoacetamide for 15 min on ice. Cells were lysed and subjected to immunoprecipitation with GFP-Trap beads according to the manufacturer's instructions (Chromotek). The immunoprecipitated integrin was released from the beads by boiling in nonreducing Laemmli buffer. Internalized integrin was blotted with anti-biotin-HRP, and the total immunoprecipitated integrin was blotted with anti-GFP antibody and used for normalization.

Statistical analysis

Statistical significance was analysed using Student's t-test for normal distribution and unequal variance or Mann Whitney test. $P < 0.05 = *$, $P < 0.01 = **$, $P < 0.005 = ***$. No statistical method was used to predetermine sample size.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

M. Humphries (University of Manchester), R. Horwitz (University of Virginia), A. Helenius (Swiss Federal Institute of Technology, ETH), M. Davidson (Florida State University) and M. Robinson (Cambridge University for Medical Research, CIMR) are acknowledged for plasmids and S. Miller (CIMR) for assistance with ITC. We thank J. Siivonen, and P. Laasola for excellent technical assistance, H. Baghirov for help with the image analysis. Turku Centre for Biotechnology Imaging Core facility and M. Saari for their help with the imaging and H. Hamidi for scientific writing and editing of the manuscript. We gratefully acknowledge the following funding sources: N.d.F. FinPharma Doctoral Program, Instrumentarium Foundation, Orion Research Foundation, Liv och Hals foundation, Finsk-Norska Medicinska Stiftelsen and the Magnus Ehrnrooth Foundation; J.I. Academy of Finland CoE, European Research Council Consolidator Grant, the Sigrid Juselius Foundation, The Finnish Heart Foundation and Finnish Cancer Organizations. D.J.O., A.G.W and T.W. are funded by Wellcome Trust fellowship 090909 (D.J.O.).

References

1. Campbell ID, Humphries MJ. Integrin structure, activation, and interactions. *Cold Spring Harb Perspect Biol.* 2011; 3(3)doi: 10.1101/cshperspect.a004994
2. Wickstrom SA, Radovanac K, Fassler R. Genetic analyses of integrin signaling. *Cold Spring Harb Perspect Biol.* 2011; 3(2)doi: 10.1101/cshperspect.a005116
3. Humphries JD, Byron A, Humphries MJ. Integrin ligands at a glance. *J Cell Sci.* 2006; 119(Pt 19): 3901–3903. [PubMed: 16988024]
4. Hemler ME, Kassner PD, Chan BM. Functional roles for integrin alpha subunit cytoplasmic domains. *Cold Spring Harb Symp Quant Biol.* 1992; 57:213–220. [PubMed: 1339660]
5. Ivaska J, Reunanen H, Westermarck J, Koivisto L, Kahari VM, Heino J. Integrin alpha2beta1 mediates isoform-specific activation of p38 and upregulation of collagen gene transcription by a mechanism involving the alpha2 cytoplasmic tail. *J Cell Biol.* 1999; 147(2):401–416. [PubMed: 10525544]

6. Mattila E, Pellinen T, Nevo J, Vuoriluoto K, Arjonen A, Ivaska J. Negative regulation of EGFR signalling through integrin- α 1 β 1-mediated activation of protein tyrosine phosphatase TCPTP. *Nat Cell Biol.* 2005; 7(1):78–85. [PubMed: 15592458]
7. Arjonen A, Alanko J, Veltel S, Ivaska J. Distinct recycling of active and inactive β 1 integrins. *Traffic.* 2012; 13(4):610–625. [PubMed: 22222055]
8. De Franceschi N, Hamidi H, Alanko J, Sahgal P, Ivaska J. Integrin traffic - the update. *J Cell Sci.* 2015; 128(5):839–852. [PubMed: 25663697]
9. Ezratty EJ, Bertaux C, Marcantonio EE, Gundersen GG. Clathrin mediates integrin endocytosis for focal adhesion disassembly in migrating cells. *J Cell Biol.* 2009; 187(5):733–747. [PubMed: 19951918]
10. Caswell PT, Vadrevu S, Norman JC. Integrins: Masters and slaves of endocytic transport. *Nat Rev Mol Cell Biol.* 2009; 10(12):843–853. [PubMed: 19904298]
11. Pellinen T, Ivaska J. Integrin traffic. *J Cell Sci.* 2006; 119(Pt 18):3723–3731. [PubMed: 16959902]
12. Mai A, Veltel S, Pellinen T, et al. Competitive binding of Rab21 and p120RasGAP to integrins regulates receptor traffic and migration. *J Cell Biol.* 2011; 194(2):291–306. [PubMed: 21768288]
13. Caswell PT, Spence HJ, Parsons M, et al. Rab25 associates with α 5 β 1 integrin to promote invasive migration in 3D microenvironments. *Dev Cell.* 2007; 13(4):496–510. [PubMed: 17925226]
14. Traub LM. Tickets to ride: Selecting cargo for clathrin-regulated internalization. *Nat Rev Mol Cell Biol.* 2009; 10(9):583–596. [PubMed: 19696796]
15. Royle SJ, Qureshi OS, Bobanovic LK, Evans PR, Owen DJ, Murrell-Lagnado RD. Non-canonical YXXGPhi endocytic motifs: Recognition by AP2 and preferential utilization in P2X4 receptors. *J Cell Sci.* 2005; 118(Pt 14):3073–3080. [PubMed: 15985462]
16. De Franceschi N, Wild K, Schlacht A, Dacks JB, Sinning I, Filippini F. Longin and GAF domains: Structural evolution and adaptation to the subcellular trafficking machinery. *Traffic.* 2014; 15(1):104–121. [PubMed: 24107188]
17. Chouhan BS, Kapyla J, Denessiouk K, Denesyuk A, Heino J, Johnson MS. Early chordate origin of the vertebrate integrin α I domains. *PLoS One.* 2014; 9(11):e112064. [PubMed: 25409021]
18. Johnson MS, Lu N, Denessiouk K, Heino J, Gullberg D. Integrins during evolution: Evolutionary trees and model organisms. *Biochim Biophys Acta.* 2009; 1788(4):779–789. [PubMed: 19161977]
19. Hedges SB. The origin and evolution of model organisms. *Nat Rev Genet.* 2002; 3(11):838–849. [PubMed: 12415314]
20. Bouvard D, Pouwels J, De Franceschi N, Ivaska J. Integrin inactivators: Balancing cellular functions in vitro and in vivo. *Nat Rev Mol Cell Biol.* 2013; 14(7):430–442. [PubMed: 23719537]
21. Honing S, Ricotta D, Krauss M, et al. Phosphatidylinositol-(4,5)-bisphosphate regulates sorting signal recognition by the clathrin-associated adaptor complex AP2. *Mol Cell.* 2005; 18(5):519–531. [PubMed: 15916959]
22. Bhunia A, Tang XY, Mohanram H, Tan SM, Bhattacharjya S. NMR solution conformations and interactions of integrin α L β 2 cytoplasmic tails. *J Biol Chem.* 2009; 284(6):3873–3884. [PubMed: 19073598]
23. Liu J, Das M, Yang J, et al. Structural mechanism of integrin inactivation by filamin. *Nat Struct Mol Biol.* 2015; 22(5):383–389. [PubMed: 25849143]
24. Metcalf DG, Moore DT, Wu Y, et al. NMR analysis of the α IIb β 3 cytoplasmic interaction suggests a mechanism for integrin regulation. *Proc Natl Acad Sci U S A.* 2010; 107(52):22481–22486. [PubMed: 21156831]
25. Yang J, Ma YQ, Page RC, Misra S, Plow EF, Qin J. Structure of an integrin α IIb β 3 transmembrane-cytoplasmic heterocomplex provides insight into integrin activation. *Proc Natl Acad Sci U S A.* 2009; 106(42):17729–17734. [PubMed: 19805198]
26. Bonet R, Vakonakis I, Campbell ID. Characterization of 14-3-3-zeta interactions with integrin tails. *J Mol Biol.* 2013; 425(17):3060–3072. [PubMed: 23763993]
27. Bridgewater RE, Norman JC, Caswell PT. Integrin trafficking at a glance. *J Cell Sci.* 2012; 125(Pt 16):3695–3701. [PubMed: 23027580]

28. Aguet F, Antonescu CN, Mettlen M, Schmid SL, Danuser G. Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. *Dev Cell*. 2013; 26(3):279–291. [PubMed: 23891661]
29. Pellinen T, Arjonen A, Vuoriluoto K, Kallio K, Fransén JA, Ivaska J. Small GTPase Rab21 regulates cell adhesion and controls endosomal traffic of beta1-integrins. *J Cell Biol*. 2006; 173(5):767–780. [PubMed: 16754960]
30. Motley AM, Berg N, Taylor MJ, et al. Functional analysis of AP-2 alpha and mu2 subunits. *Mol Biol Cell*. 2006; 17(12):5298–5308. [PubMed: 17035630]
31. Somasundaram B, Norman JC, Mahaut-Smith MP. Primaquine, an inhibitor of vesicular transport, blocks the calcium-release-activated current in rat megakaryocytes. *Biochem J*. 1995; 309(Pt 3):725–729. [PubMed: 7639685]
32. Nishimura T, Kaibuchi K. Numb controls integrin endocytosis for directional cell migration with aPKC and PAR-3. *Dev Cell*. 2007; 13(1):15–28. [PubMed: 17609107]
33. Teckchandani A, Mulkearns EE, Randolph TW, Toida N, Cooper JA. The clathrin adaptor Dab2 recruits EH domain scaffold proteins to regulate integrin beta1 endocytosis. *Mol Biol Cell*. 2012; 23(15):2905–2916. [PubMed: 22648170]
34. Sandri C, Caccavari F, Valdembri D, et al. The R-ras/RIN2/Rab5 complex controls endothelial cell adhesion and morphogenesis via active integrin endocytosis and rac signaling. *Cell Res*. 2012; 22(10):1479–1501. [PubMed: 22825554]
35. Pellinen T, Tuomi S, Arjonen A, et al. Integrin trafficking regulated by Rab21 is necessary for cytokinesis. *Dev Cell*. 2008; 15(3):371–385. [PubMed: 18804435]
36. Wennerberg K, Lohikangas L, Gullberg D, Pfaff M, Johansson S, Fassler R. Beta 1 integrin-dependent and -independent polymerization of fibronectin. *J Cell Biol*. 1996; 132(1-2):227–238. [PubMed: 8567726]
37. Mai A, Muharram G, Barrow-McGee R, et al. Distinct c-met activation mechanisms induce cell rounding or invasion through pathways involving integrins, RhoA and HIP1. *J Cell Sci*. 2014; 127(Pt 9):1938–1952. [PubMed: 24790222]
38. Chao WT, Kunz J. Focal adhesion disassembly requires clathrin-dependent endocytosis of integrins. *FEBS Lett*. 2009; 583(8):1337–1343. [PubMed: 19306879]
39. Hirst J, Irving C, Borner GH. Adaptor protein complexes AP-4 and AP-5: New players in endosomal trafficking and progressive spastic paraplegia. *Traffic*. 2013; 14(2):153–164. [PubMed: 23167973]
40. Liu L, He B, Liu WM, Zhou D, Cox JV, Zhang XA. Tetraspanin CD151 promotes cell migration by regulating integrin trafficking. *J Biol Chem*. 2007; 282(43):31631–31642. [PubMed: 17716972]
41. Morse EM, Brahme NN, Calderwood DA. Integrin cytoplasmic tail interactions. *Biochemistry*. 2014; 53(5):810–820. [PubMed: 24467163]
42. Ivaska J, Heino J. Cooperation between integrins and growth factor receptors in signaling and endocytosis. *Annu Rev Cell Dev Biol*. 2011; 27:291–320. [PubMed: 21663443]
43. Zaidel-Bar R, Geiger B. The switchable integrin adhesome. *J Cell Sci*. 2010; 123(Pt 9):1385–1388. [PubMed: 20410370]
44. Calderwood DA, Fujioka Y, de Pereda JM, et al. Integrin beta cytoplasmic domain interactions with phosphotyrosine-binding domains: A structural prototype for diversity in integrin signaling. *Proc Natl Acad Sci U S A*. 2003; 100(5):2272–2277. [PubMed: 12606711]
45. Santolini E, Puri C, Salcini AE, et al. Numb is an endocytic protein. *J Cell Biol*. 2000; 151(6):1345–1352. [PubMed: 11121447]
46. Mishra SK, Hawryluk MJ, Brett TJ, et al. Dual engagement regulation of protein interactions with the AP-2 adaptor alpha appendage. *J Biol Chem*. 2004; 279(44):46191–46203. [PubMed: 15292237]

Methods references

47. Marchler-Bauer A, Derbyshire MK, Gonzales NR, et al. CDD: NCBI's conserved domain database. *Nucleic Acids Res*. 2015; 43(Database issue):D222–6. [PubMed: 25414356]

48. Finn RD, Bateman A, Clements J, et al. Pfam: The protein families database. *Nucleic Acids Res.* 2014; 42(Database issue):D222–30. [PubMed: 24288371]
49. Curwen V, Eyraas E, Andrews TD, et al. The ensembl automatic gene annotation system. *Genome Res.* 2004; 14(5):942–950. [PubMed: 15123590]
50. Owen DJ, Evans PR. A structural explanation for the recognition of tyrosine-based endocytotic signals. *Science.* 1998; 282(5392):1327–1332. [PubMed: 9812899]
51. Parsons M, Messent AJ, Humphries JD, Deakin NO, Humphries MJ. Quantification of integrin receptor agonism by fluorescence lifetime imaging. *J Cell Sci.* 2008; 121(Pt 3):265–271. [PubMed: 18216331]
52. Laukaitis CM, Webb DJ, Donais K, Horwitz AF. Differential dynamics of alpha 5 integrin, paxillin, and alpha-actinin during formation and disassembly of adhesions in migrating cells. *J Cell Biol.* 2001; 153(7):1427–1440. [PubMed: 11425873]
53. Vonderheit A, Helenius A. Rab7 associates with early endosomes to mediate sorting and transport of semliki forest virus to late endosomes. *PLoS Biol.* 2005; 3(7):e233. [PubMed: 15954801]
54. Virtakoivu R, Mai A, Mattila E, et al. Vimentin-ERK signaling uncouples slug gene regulatory function. *Cancer Res.* 2015; 75(11):2349–2362. [PubMed: 25855378]
55. Gu Z, Noss EH, Hsu VW, Brenner MB. Integrins traffic rapidly via circular dorsal ruffles and macropinocytosis during stimulated cell migration. *J Cell Biol.* 2011; 193(1):61–70. [PubMed: 21464228]

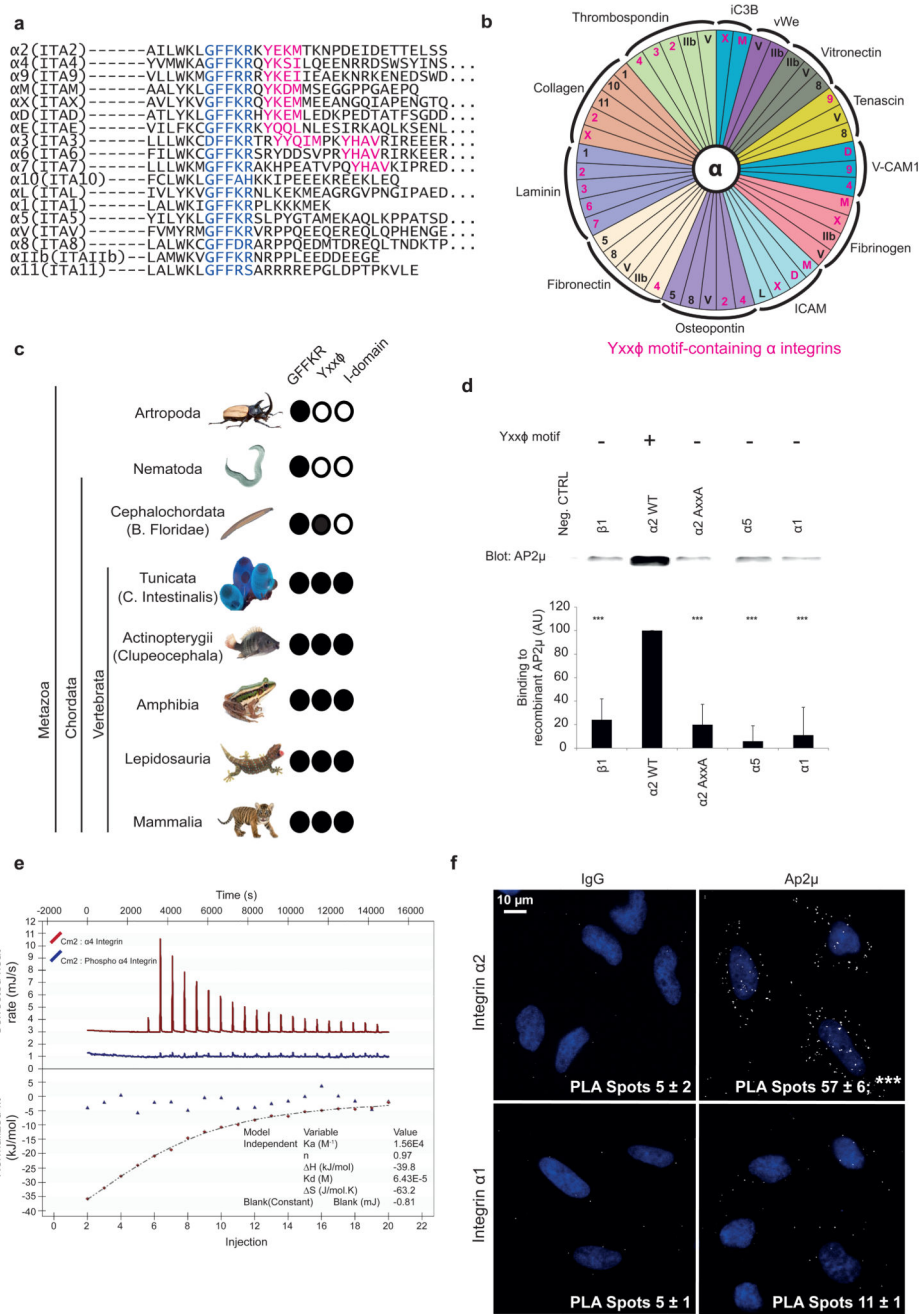


Figure 1. Evolutionarily conserved Yxxφ motif defines a subset of integrin α-chains. (a) Alignment of cytoplasmic regions of all α integrins from *Homo sapiens*. Blue, conserved membrane-proximal GFFKR region; pink, Yxxφ motif. (b) Distribution of the Yxxφ motif in relation to ligand binding specificity of integrins. Pink, α-integrin subunits containing the Yxxφ motif. (c) Taxonomic distribution of the Yxxφ motif across Metazoa. Distribution of the membrane-proximal GFFKR motif and of the I domain are shown for comparison. Images reproduced with permission from Bob Goldstein (University of North Carolina at Chapel

Hill) under the terms of the CC BY-SA 3.0 license (worm); from Phiseksit (beetle), khunaspix (fish), Kamnuan (lizard), japanachai (frog) and anankml (tiger) at <http://www.freedigitalphotos.net>; and from <http://www.istockphoto.com> (tunicate). Image of lancelet is courtesy of M. Laasola. **(d)** Pulldown assay with recombinant C- μ 2 and biotinylated α -integrin peptides. Neg. ctrl, beads alone. Quantification shows binding (with neg. ctrl subtracted) relative to α 2 WT. Bars indicate mean \pm s.e.m. ($n = 7$ independent binding assays). *** $P < 0.001$ (unpaired Student's t test; two-tail distribution). Uncropped blot is presented in Supplementary Data Set 1. AU, arbitrary units. **(e)** Representative isothermal titration calorimetry of integrin- α 4 peptide binding to C- μ 2 (red). The experiment was performed on six occasions on two different protein preparations, yielding a final overall estimate for the K_d for the interaction of $72 \pm 6 \mu\text{M}$ with $n = 1.0 \pm 0.1$ (mean \pm s.e.m.). Blue line indicates lack of detectable binding for the same peptide when the tyrosine residue was phosphorylated (phospho- α 4 integrin). **(f)** Proximity ligation assay between endogenous AP2 and endogenous α 1 and α 2 integrins. Numbers indicate mean \pm s.e.m. ($n = 3$ independent cell cultures). *** $P < 0.001$ (two-tailed Mann-Whitney test).

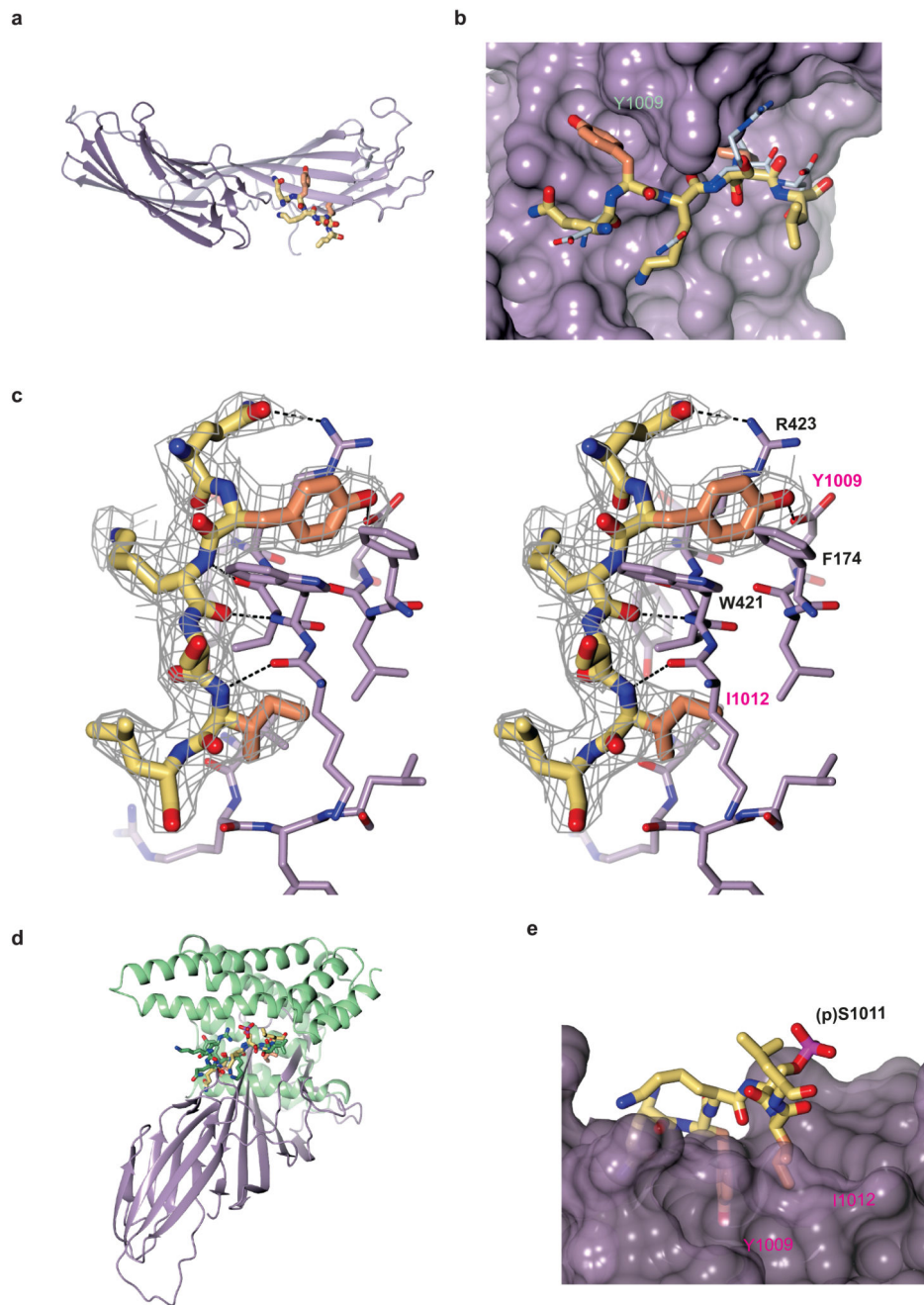
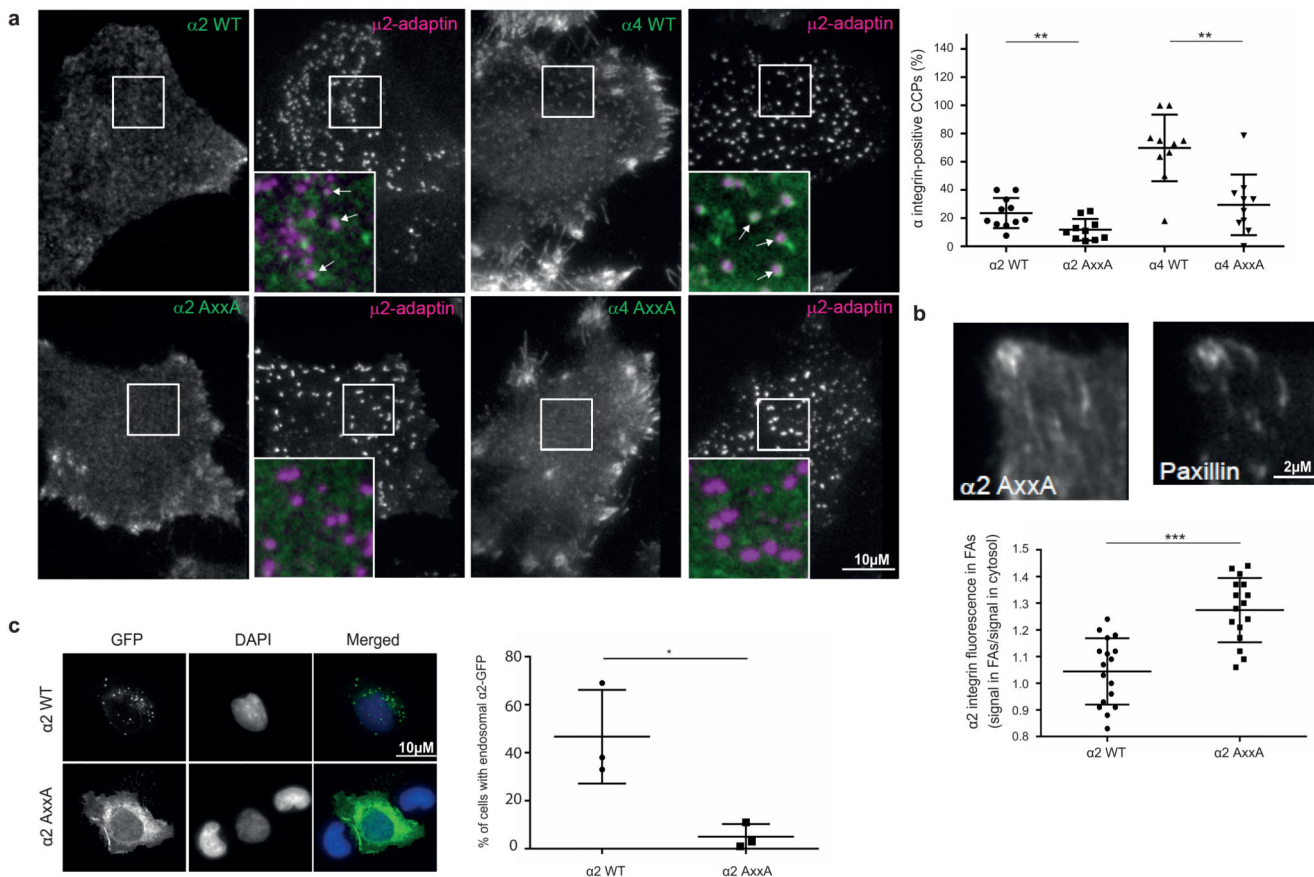


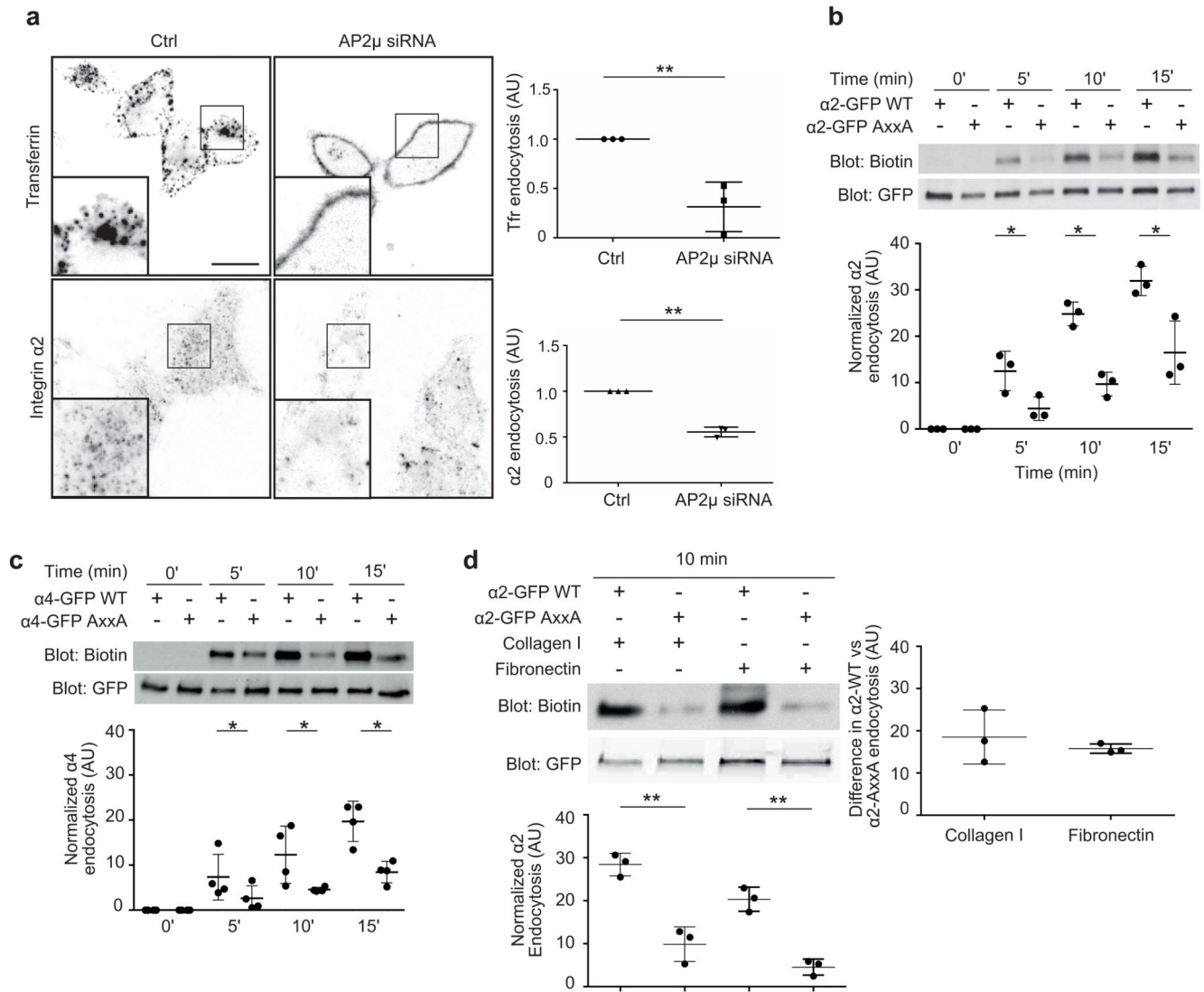
Figure 2.

Structural analysis of the integrin $\alpha 4$ -AP2 C- $\mu 2$ complex. (a) Structure of the $\alpha 4$ -integrin cytoplasmic domain in complex with C- $\mu 2$. C- $\mu 2$ is shown with subdomain 1 in pale purple and subdomain 2 in dark purple. The residues in the YxxI motif critical for the interaction are indicated in orange within the rest of the $\alpha 4$ -integrin peptide shown in yellow. (b) Magnified view of $\alpha 4$ -integrin peptide binding in the same view as in a, shown on a surface representation of C- $\mu 2$. Carbon atoms in the residues critical for binding are colored orange; those from other residues are colored yellow, with oxygen and nitrogen atoms colored red

and blue. The position where the YxxΦ motif of TGN38 (PDB 1BXX) would bind (white) is shown for comparison, as determined by overlaying the C-μ2 domains of the two structures. (c) Stereo representation of the binding of the α4-integrin YKSL (carbons in yellow and orange) to C-μ2 (carbons in lilac). Electron density shown (gray) is an omit map contoured at 0.09 e⁻/Å³. (d) Superposition of α4 and phosphoserine α4-integrin peptides in complex with C-μ2 (purple) and 14-3-3ζ (green), respectively. (e) Modeling of integrin-α4 phosphopeptide (QYK(p)SIL) in the C-μ2 binding pocket. (p) denotes phosphorylation.

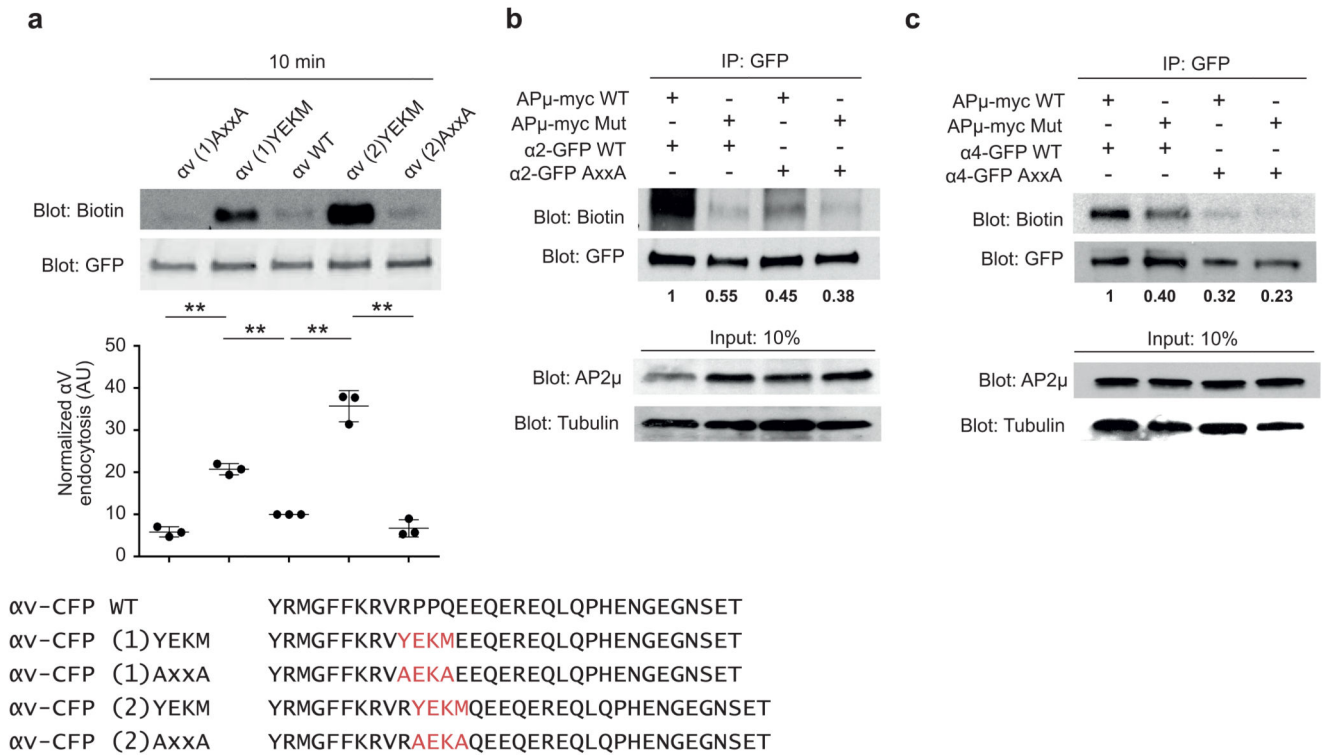
**Figure 3.**

YxxΦ motif controls integrin recruitment to CCPs. **(a)** Colocalization between GFP-tagged α2 or α4 integrins (WT or AxxA mutants) and μ2-adaptin-mCherry. Left, the TIRF plane is shown. Arrows indicate colocalized puncta. Scale bar, 10 μM. Right, quantification of percentage of μ2-adaptin-mCherry CCPs positive for integrin-GFP. Each spot represents the measurement for one cell; bars indicate mean ± s.d. ($n = 10$ cells per condition; data are from two experiments with independent cell cultures). ** $P < 0.01$ (two-tailed Mann-Whitney test). **(b)** Recruitment of GFP-tagged α2-integrin WT or AxxA mutant to focal adhesions (FAs), as visualized by paxillin staining. The TIRF plane is shown. Scale bar, 2 μM. Quantification of α2-integrin GFP signal intensity in focal adhesions compared to plasma membrane outside of focal adhesions. Each dot represents the value for one cell; bars indicate mean ± s.d. ($n = 16$ cells per condition; data are from two experiments with independent cell cultures). *** $P < 0.001$ (two-tailed Mann-Whitney test). **(c)** Localization of GFP-tagged α2-integrin WT or AxxA in HEK293 cells. Maximum intensity projection of z stack from confocal microscopy is shown. Scale bar, 10 μm. Quantification shows mean ± s.e.m. ($n = 3$ independent cell cultures). * $P < 0.05$ (unpaired Student's t test; two-tail distribution).

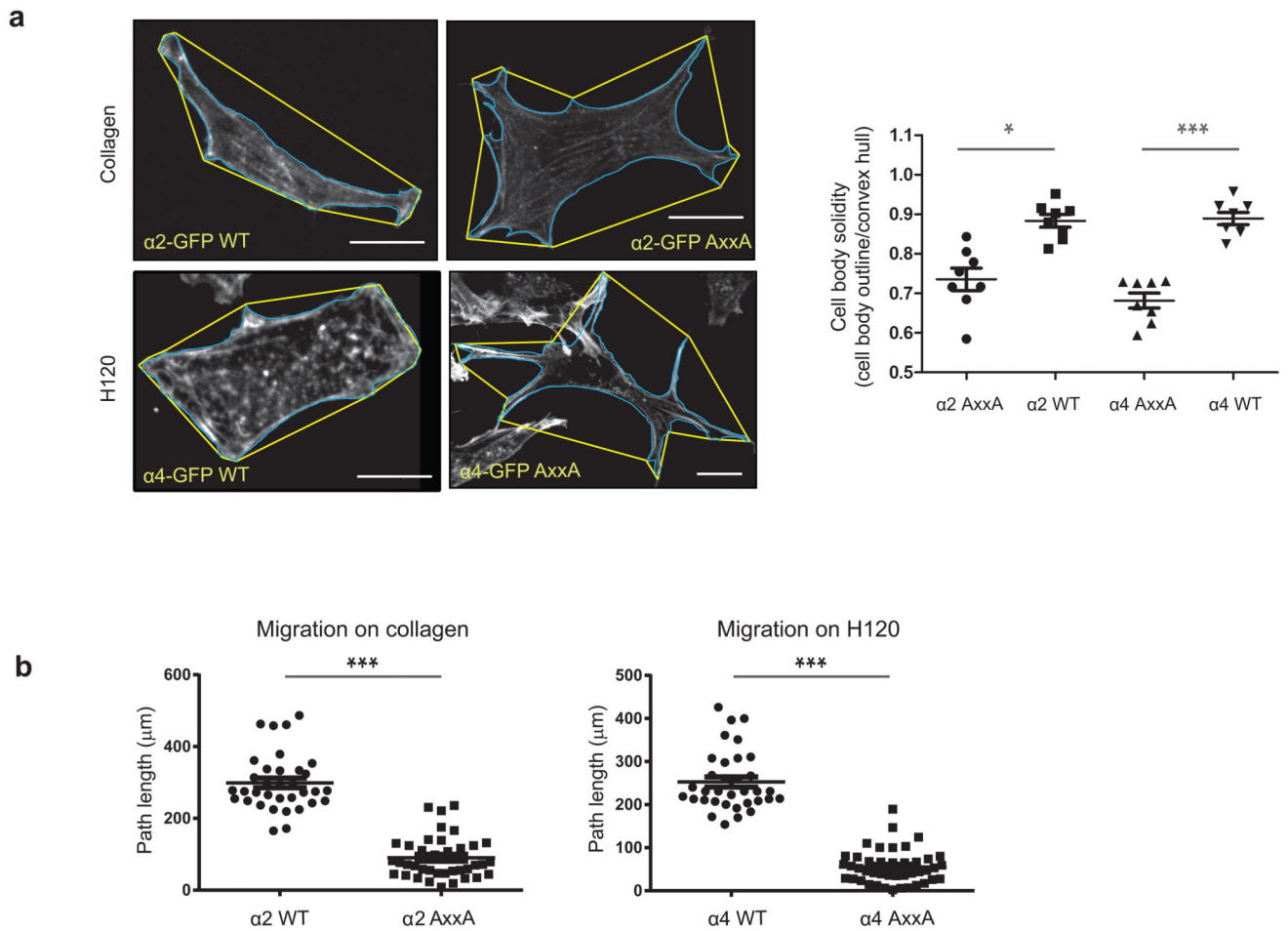
**Figure 4.**

Yxx ϕ motif regulates integrin $\alpha 2$ and $\alpha 4$ endocytosis. (a) Endocytosis of antibody-labeled cell-surface endogenous $\alpha 2$ -integrin and transferrin receptor in HeLa cells silenced with control (ctrl) or AP2 μ siRNA. The signal of the internalized antibody was analyzed after 30 min endocytosis. Quantification shows mean \pm s.e.m. ($n = 3$ independent cell cultures). In total, 120 cells were analyzed for transferrin receptor and 300 cells for $\alpha 2$ integrin. $**P < 0.01$ (unpaired Student's t test; two-tail distribution). (b,c) Biotin-based endocytosis assays in HEK293 cells expressing either $\alpha 2$ -GFP (b) or $\alpha 4$ -GFP (c) WT or AxxA mutant. Endocytosis was allowed for the indicated times, and biotin signal was normalized against total $\alpha 2$ - and $\alpha 4$ -GFP amounts measured from the GFP blot. Bars indicate mean \pm s.e.m. ($n = 3$ ($\alpha 2$) and $n = 4$ ($\alpha 4$) independent cell cultures). $*P < 0.05$ (unpaired Student's t test; two-tail distribution). Uncropped blots are presented in Supplementary Data Set 1. (d) Biotin-based endocytosis assays in HEK293 cells expressing either GFP-tagged $\alpha 2$ -integrin WT or AxxA mutant plated on collagen I- or fibronectin- (5 μ g/ml) coated dishes. Endocytosis was

allowed for 10 min, and biotin signal was normalized against the total $\alpha 2$ -GFP amount measured from the GFP blot. Bars indicate mean \pm s.e.m. ($n = 3$ independent cell cultures). ** $P < 0.01$ (unpaired Student's t test; two-tail distribution). Right, difference between endocytosed WT and mutant integrin on collagen and fibronectin (endocytosed $\alpha 2$ WT with $\alpha 2$ AxxA subtracted). Uncropped blots are presented in Supplementary Data Set 1.

**Figure 5.**

Introduction of the Yxx ϕ motif induces endocytosis of integrin α_V . **(a)** Biotin-based endocytosis assays in HEK293 cells expressing CFP-tagged α_V -integrin WT, YEKM GOF mutant or AxxA loss-of-function mutant. Sequences shown at the bottom. Endocytosis was allowed for 10 min, and biotin signal was normalized against the total α_V amount measured from the GFP blot. IP, immunoprecipitation. Bars indicate mean \pm s.e.m. ($n = 3$ independent cell cultures). ** $P < 0.01$ (unpaired Student's t test; two-tail distribution). Uncropped blots are presented in Supplementary Data Set 1. **(b,c)** Biotin-based endocytosis assays in HEK293 cells stably expressing AP2 μ WT or F174A D176S mutant, and transiently expressing α_2 -GFP **(b)** or α_4 -GFP **(c)** WT or AxxA mutant. β -tubulin is a loading control. Biotin signal was normalized against the total α_2 - and α_4 -GTP amounts measured from the GFP blot. Time point, 15 min. Numbers indicate mean from 2 experiments (with independent cell cultures). Uncropped blots are presented in Supplementary Data Set 1.

**Figure 6.**

Yxx ϕ mutagenesis impairs cell migration mediated by $\alpha 2$ and $\alpha 4$ integrin. **(a)** Cell shape analysis of phalloidin-stained GD25 β 1A cells expressing either $\alpha 2$ -GFP or $\alpha 4$ -GFP WT or AxxA mutant and adhering to collagen ($\alpha 2$) or fibronectin fragment H120 ($\alpha 4$). Solidity was calculated as cell-body outline (blue) divided by convex hull (yellow). Each dot represents one cell; bars indicate mean \pm s.e.m. ($n = 8$ cells per condition; data are from two experiments with independent cell cultures) * $P < 0.05$; *** $P < 0.001$ (two-tailed Mann-Whitney test). Scale bars, 10 μm . **(b)** Quantification of migration of GD25 β 1A cells expressing either $\alpha 2$ -GFP or $\alpha 4$ -GFP WT or AxxA mutant and adhering to collagen ($\alpha 2$) or fibronectin fragment H120 ($\alpha 4$). Cells were imaged for 16 h, and the path length of the tracks was analyzed. In total, 35 cells for each condition were analyzed, from 3 independent cell cultures. *** $P < 0.001$ (two-tailed Mann-Whitney test).

Table 1
Data collection and refinement statistics

C-μ2-α₄ peptide	
Data collection	
Space group	<i>P</i> 6 ₄
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	42.08, 86.93, 86.24
α , β , γ (°)	118.11, 99.65, 95.47
Resolution (Å)	2.75 (2.90 - 2.75) ^a
<i>R</i> _{meas}	0.078 (0.415)
<i>I</i> / σ <i>I</i>	14.2 (3.7)
<i>CC</i> _{1/2} ^b	0.997 (0.776)
Completeness (%)	97.2 (80.8)
Redundancy	4.9 (4.6)
Refinement	
Resolution (Å)	2.77
No. reflections	16115
<i>R</i> _{work} / <i>R</i> _{free}	0.174 / 0.224
No. atoms	
Protein	2130
Water	45
<i>B</i> factors	
Protein	62
Water	48
r.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	2.1

Data were collected from one crystal

^aValues in parentheses are for highest-resolution shell.

^b*CC*_{1/2} is the correlation coefficient between random half-data sets.