The N-terminal domain of the R28 protein promotes emm28 group A Streptococcus adhesion to host cells via direct binding to three integrins

Antonin Weckel, Dorian Ahamada, Samuel Bellais, Céline Méhats, Céline Plainvert, Magalie Longo, Claire Poyart, Agnès Fouet

The N-terminal domain of the R28 protein promotes emm28 group A Streptococcus adhesion to host cells via direct binding to three integrins

Antonin Weckel[1,2,3], Dorian Ahamada,[1,2,3], Samuel Bellais[1,2,3,†], Céline Méhats[1,2,3], Céline Plainvert[1,2,3,4,5], Magalie Longo[1,2,3], Claire Poyart[1,2,3,4,5], Agnès Fouet[1,2,3,4*]

From [1]INSERM U1016, Institut Cochin, [2]CNRS UMR 8104, [3]Université Paris Descartes (UMR-S1016) Paris France, [4]Centre Nationale de Référence des Streptocoques, [5]Hôpitaux Universitaires Paris Centre, Cochin, Assistance Publique Hôpitaux de Paris, France

Running title**:** GAS R28_{Nt} interacts with integrins to promote adhesion

* To whom correspondence should be addressed: Institut Cochin, 22 rue Méchain 75014 Paris, France; agnes.fouet@inserm.fr; Tel: + 33 1 40 51 64 50; Fax 33: + 1 40 51 64 54

† present address: Bioaster, 28 rue du Dr Roux, 75015 Paris

ABSTRACT

Group A Streptococcus (GAS) is a human-specific pathogen responsible for a wide range of diseases, from superficial to life-threatening invasive infections, including endometritis, and auto-immune sequelae. GAS strains express a vast repertoire of virulence factors that varies depending on the strain genotype and many adhesins are restricted to some genotypes. The *emm28* genotype is the third most prevalent for invasive infections in France and is associated with gyneco-obstetrical infections. R28, a cell wall-anchored surface protein harbored by *emm28* strains, was previously described to promote adhesion to cervical epithelial cells. This study aimed to determine by cellular and biochemical approaches whether R28 promotes adhesion to other cells and at characterizing its receptor. We show that through its N-terminal domain, R28_{Nt}, R28 promotes bacterial adhesion to endometrial epithelial and stromal cells. R28_{Nt} has been further subdivided into two domains and both are involved in cell binding. R28_{Nt} and both subdomains interact directly with the laminin-binding α3β1, α6β1 and α6β4 integrins; interestingly, these bindings do not require the presence of divalent cations. R28 is the first GAS adhesin described to bind directly to these integrins that are expressed by most epithelial cells. Finally, R28_{Nt} also promotes binding to keratinocytes and pulmonary epithelial cells, suggesting that it may be involved in the prevalence of the *emm28* genotype.

Streptococcus pyogenes, also known as Group A streptococcus (GAS), is a Gram-positive bacterium responsible for a wide range of diseases, from superficial infections such as pharyngitis and dermatitis, to severe invasive infections such as necrotizing fasciitis and endometritis (1–3). GAS infections are also responsible for post-infectious complications such as rheumatic arthritis and glomerulonephritis and, altogether, GAS infections are responsible for 517,000 deaths annually worldwide (4).

GAS strains are genetically diverse and are genotyped through sequencing of the 5' end of the *emm* gene (5) encoding the M protein, a major virulence factor; more than 250 *emm*-types have been described (6). A link exists between genotype and tissue tropism, with throat- and skin-specialists and ubiquitous genotypes (7); a link between genotype of invasive strains and elicited disease exists for some but not all genotypes (our unpublished data, https://cnr-strep.fr/). Approximately 10 % of the GAS genome is composed of exogenous genetic elements encoding virulence factors, with substantial variation between different *emm*-types, which could account for the tropism (8). The first to third most prevalent genotype in Europe, *emm28*, is responsible for 8% of GAS invasive infections in France (9, 10, https://cnr-strep.fr/). It is associated to endometritis and for example, in France, 27% of GAS invasive infections in women occur in the gyneco-obstetrical sphere (10–13). *emm28* strains harbor an integrative conjugative element named RD2

that was likely horizontally transferred from *Streptococcus agalactiae*, also known as Group B Streptococcus (GBS) (13, 14). GBS colonizes 10-30% of healthy women's urogenital-tract (15) and it was suggested that the presence of this integrative conjugative element accounts for the *emm28* GAS gyneco-obstetrical tropism (13). This remarkable tissue association together with the high prevalence of *emm*28 strain invasive infections prompted us to study the role of R28, an RD2 encoded surface protein, in GAS *emm*28 infections.

Adhesion to host tissues is the initial step for all GAS infections. It is mediated by different factors, mainly surface proteins, that bind either extracellular matrix components, indirectly the cell surface through plasmatic or extracellular matrix components bridging, or directly to eukaryotic receptors [reviewed in (16) and (17)]. RD2 encodes four surface proteins including R28 and a R28-deficient *emm28* strain is non-adherent to the cervical cell line ME180 (18). R28 is a member of the Alp family composed of GBS proteins that share evolutionary and structural similarity and which includes the alpha C protein (also known as ACP or α), Rib, R28 (also known as Alp3 in GBS) and Alp2 (18–21) (Fig. 1A). These chimeric proteins are composed of a signal peptide, an N-terminal domain, repeats, and an LPXTG anchoring motif. The repeats number varies among clinical isolates and their structure is related to the Ig-like fold (22). The repeats are identical within one protein and the identity percentage between Alp members varies between 35 and 94, the latter being between Rib and R28 repeats. Repeats are considered to properly expose the N-terminal domain at the bacterial surface, potentially the functional domain (21). This domain is composed of one, for ACP and Rib, two, for R28, and three modules, for Alp2. The first modules of R28 and Alp2 share 99% identity and are 70 and 56% identical to the module of ACP and Rib, respectively. The second modules of R28 and Alp2 are 99% identical and is similar to the β protein, unrelated to the Alp family. Alp2 third module is a repeat of the second one (Fig. 1A) (21, 23, 24). Thus, R28 is a chimer of ACP and the β protein for its N-terminal domain and Rib for its repeats and it immunologically cross-reacts with both proteins (19, 25). ACP is the most studied member of the Alp family and the crystal structure of the N-terminal domain has been elucidated (24). It binds glycosaminoglycans in a region encompassing the end of the N-terminal domain and the repeats (26,

27), and the α1β1 integrin through a KTD motif localized in a β-sandwich subdomain at the N-terminal end of its N-terminal domain (24, 28, 29); this motif is absent from R28, Alp2 and Rib (Fig. 1A). These interactions increase GBS internalization in the cervical cell line ME180 (29).

In this study, we analyzed the capacity of R28 to promote adhesion to host cells and characterized the adhesion domain. We demonstrate that the N-terminal domain of R28 (R28$_{Nt}$) is sufficient to promote direct adhesion to different gyneco-obstetrical cell lines, and further identified two subdomains within R28$_{Nt}$ that are both involved in this adhesion process. We then characterized the chemical nature of the R28$_{Nt}$ receptor and isolated different ligands. We show that R28$_{Nt}$ interacts with the laminin-binding integrins α3β1, α6β1 and α6β4. Finally, we show that R28$_{Nt}$ increases adhesion also to skin and pulmonary cells, further extending the role of R28 as a GAS adhesin involved in GAS *emm*28 prevalence.

Results
The N-terminal domain of R28 is sufficient to promote adhesion to human female genital tract cells

Association of GAS *emm28* strains with gyneco-obstetrical infections could be a consequence of its capacity to colonize the vaginal tract (10). A GAS R28-deleted mutant adheres less than the parental strain to cells from the cervical ME180 lineage (18). To assay the role of R28 in a more physiological situation, we tested whether the phenotype could also be observed on human decidual stromal cells (hDSC) isolated from decidual biopsy of specimens obtained after caesarian delivery (Fig. 2A). The decidua is the lining of the uterine during pregnancy and consists of differentiated endometrial stromal fibroblasts and recruited leukocytes. It may be a direct target of GAS infection, a major cause of severe puerperal sepsis. The R28-deleted strain adhered significantly less (18%) to hDSC than the wild-type strain (Fig. 2A), confirming the role of R28 as an adhesin on physiologically relevant cells. We then sought to characterize the adhesion domain of the R28 protein using different cells from the female genital tract, hDSCs as well as two cell lineages, cervical cells (ME180), often used to study female genital tract infections, and endometrial epithelial cells (HEC-1-A). The six to 17 highly conserved 79-residue long repeats of

R28 are considered to expose the 368 residue-long N-terminal domain at the bacterial surface, potentially the functional domain (Fig. 1A**)** (18, 29). Consequently, to analyze the ability of R28 to promote adhesion to various cells, we focused on the N-terminal domain of R28, R28$_{Nt}$, using biotinylated R28$_{Nt}$ (Fig. 2B-D). R28$_{Nt}$ displays higher binding capacity to hDSC, ME180 and HEC-1-A cells than soluble BSA (200, 10 and 10 times more binding at 10 µM, respectively).

To confirm the capacity of R28$_{Nt}$ to promote adhesion, we tested whether R28$_{Nt}$-coated beads or a heterologous bacterium expressing R28$_{Nt}$ could bind more to HEC-1-A cells than the controls (Fig. 2E-I). R28$_{Nt}$-coated beads adhered three times more than the BSA-coated control beads (p<0.05) (Fig. 2E-G) and this binding was blocked by the addition of purified R28$_{Nt}$ antibodies (p<0.05) (Fig. 2H). Also, the *Lactococcus lactis* strain expressing R28$_{Nt}$ (Fig. S1A) adhered significantly more than the control strain harboring the empty vector (+35%, p = 0.0066) (Fig. 2H). The increased adherence may seem weak. However, in the absence of the repeats, R28$_{Nt}$ may be poorly exposed at the *L. lactis* cell surface or hidden by the protective polysaccharide pellicle (30) compared to the exposition of the complete R28 protein in GAS *emm*28 strains; this could lead to an underestimation of R28$_{Nt}$ capacity to promote adhesion. Thus, the N-terminal domain of R28 is sufficient to increase the adherence of a Gram-positive bacterium to endometrial cells.

Altogether these results demonstrate that R28$_{Nt}$ promotes bacterial adhesion to female genital tract cells.

Both N-terminal subdomains R28-N1 and R28-N2 are sufficient to promote adhesion to HEC-1-A cells.

The R28$_{Nt}$ domain is composed of two halves, from amino-acid residue 56 to 229 and from residue 230 to 424, which we termed R28-N1 and R28-N2, respectively (Fig. 1A and Fig. S1B (21). A BLAST alignment indicated that R28-N1 and R28-N2 share no similarity (E-value, 0.27). To test which sub-domain mediates adhesion, we produced and purified the corresponding peptides (Fig. S1B, S1C) and incubated them after biotinylation with ME180 and HEC-1-A cells (Fig. 3A-C). Both the R28-N1 and R28-N2 subdomains showed significant binding with both cell types compared to BSA (p<0.001 for both cell-types), with R28-N2 displaying an affinity two-to three-fold higher

than that of R28-N1 (Fig. 3C). The R28-N1 and R28-N2 binding values indicate an additive contribution of R28-N1 and R28-N2 to the R28$_{Nt}$ binding on HEC-1-A cells and a synergetic one on ME180 cells (Fig. 3C).

To assess the binding of these subdomains in more physiological conditions, we tested whether R28-N1- and R28-N2-coated beads bound HEC-1-A cells (Fig. 3D). Twenty-seven % and 19% of R28-N1- and R28-N2-coated beads bound to HEC-1-A cells, respectively, a percentage significantly higher than BSA-coated beads (6.3%, p<0.001 and p<0.05 for R28-N1 and R28-N2 with HEC-1-A cells, respectively). We checked the binding specificity of the coated beads by incubating cells with an excess amount of soluble peptides (20 µM) and assessing the percentage of beads still bound to cells (Fig. 3D). Purified R28-N1 and R28-N2 peptides competed with their respective coated beads (6.3% and 4.5 % vs 27 % and 19%, for R28-N1 and R28-N2, respectively) confirming the binding specificity. Altogether, these data clearly indicate that both the R28-N1 and R28-N2 subdomains are sufficient to promote adhesion to HEC-1-A cells.

We then sought to determine if R28-N1 and R28-N2 have independent receptors. We performed competition assays in which we assessed the percentages of R28-N1- and R28-N2-coated beads bound to cells pre-incubated with excess amounts (20 µM) of soluble R28-N2 and R28-N1, respectively (Fig. 3D). These subdomains compete with one another (p<0.005), suggesting that these subdomains share common receptor(s). Moreover, both displaced R28$_{Nt}$-coated beads lowering its binding level to unspecific binding level (BSA-coated beads, 6%), further supporting that R28-N1 and R28-N2 share a common receptor.

R28$_{Nt}$ adhesive properties are not conserved among all Alp family members

Among members of the Alp family, alpha C protein (ACP) binds, through its N-terminal domain, the α1β1 integrin (28) and, through a region encompassing the end of the N-terminal domain and the repeats, glycosaminoglycans (26, 27). In contrast, attempts to demonstrate that Rib is an adhesin have been unsuccessful (31). We focused on the shared N-terminal domains of these proteins, leaving out from ACP$_{Nt}$ the region binding glycosaminoglycans. A BLAST alignment indicated that R28-N1 is 70% and 56% identical to the N-terminal domains of ACP and Rib respectively (Fig. 1A). Furthermore, a Phyre

analysis predicted, with a confidence of 100%, that R28-N1 has the same structure as ACP, that is two domains, an N-terminal β-sandwich, sharing structural elements with the type III fibronectin fold, and a C-terminal three-helix bundle (Fig. 1B) (24). We thus wondered whether the adhesive properties of R28-N1 to cervical and endometrial cells are conserved among the ACP or Rib N-terminal domains. We tested the binding capacity of ACP$_{Nt}$ and Rib$_{Nt+2R}$ which contains, in addition to the N-terminal domain, two repeats which are 94% identity to the R28 repeats (Fig. 1A). We expressed and purified these peptides and analyzed their direct binding to ME180 and HEC-1-A cells (Fig. S1C, S1D; Fig. 3A, 3B). ACP$_{Nt}$ binds to both cell-types, yet less than R28-N1 and consequently even less than R28-N2 or R28$_{Nt}$. Since ACP$_{Nt}$ binds HEC-1-A, despite the absence of α1β1 integrin on its surface (32), we tested whether it shared a receptor with R28-N1 or R28-N2 by assaying, as for the R28-N1 and R28-N2 competitions, whether it can compete with R28-N1 and R28-N2 (Fig. 3D, 3E). ACP$_{Nt}$ competes with itself and R28-N1 (p<0.005), suggesting that these subdomains share a common receptor, but not with R28-N2 (Fig. 3B). Moreover, we did not detect a competition between R28-N1 and ACP-bound beads (Fig. 3E). In contrast to ACP, Rib$_{Nt+2R}$ does not bind to either ME180 or HEC-1-A cells, which indicates that this binding property is restricted to some Alp family members (Fig. 3 A, 3B). Moreover, because Rib$_{Nt+2R}$ contains two R28-like repeats, its inability to bind ME180 and HEC-1-A cells suggests that R28 repeats do not mediate the adhesion on their own.

R28$_{Nt}$ receptor is a membrane protein

Our experiments were carried out in the absence of added extracellular or plasmatic protein; we consequently hypothesized that R28$_{Nt}$ binds directly to a cell surface receptor(s). To identify it, we first defined its chemical nature. As previously described (33), we applied to HEC-1-A cells different treatments that affect potential receptors depending on their chemical nature before incubating them with soluble R28$_{Nt}$ and quantifying bound R28$_{Nt}$ by flow cytometry (Fig. 4). Pronase and trypsin treatments significantly reduced, in a dose dependent manner, the percentage of positively labelled cells down to 10% and 29%, respectively. Also, neither heparinase I nor sodium periodate treatments affected R28$_{Nt}$ binding to cells (data not shown), excluding an interaction between R28$_{Nt}$ and

glycosaminoglycans or carbohydrates which harbor vicinal hydroxyl groups. These results suggest that the R28$_{Nt}$ receptor is a cell-surface exposed protein. Similar results were obtained with R28-N1 and R28-N2 (Fig. S2A-S2B). Altogether, our data suggest that the R28$_{Nt}$, R28-N1 and R28-N2 receptor(s) is (are) a cell-surface exposed protein(s).

Identification of several putative receptors of R28$_{Nt}$

To identify R28 receptor(s), a co-immunoprecipitation experiment using HEC-1-A cells was set up followed by mass spectrometry analysis. Cross-linked R28$_{Nt}$ - cellular proteins complexes were resolved on acrylamide gels and Western blotting highlighted multiple high molecular, above 170 kDa, complexes (data not shown) absent in the co-immunoprecipitation without R28$_{Nt}$ incubation (control). The zones above 170 kDa of R28$_{Nt}$ and control immunoprecipitations were excised from gels and the protein contents were analyzed by mass spectrometry (Supplementary file). Hits enriched in the co-immunoprecipitation in the presence of the bait R28$_{Nt}$ in all three independent experiments were initially selected. From this list, and since the R28$_{Nt}$ receptors are cell-surface exposed proteins (Fig. 4), only such proteins were further selected (Table 1). Hence, through co-immunoprecipitation of R28$_{Nt}$ we identified several potential R28$_{Nt}$ receptors.

R28$_{Nt}$, R28-N1 and R28-N2 interact directly with integrins

Of all the proteins identified as present in the complexes co-immunoprecipitated with R28$_{Nt}$, we decided to focus on integrins as possible R28$_{Nt}$ receptors. To test this, we assayed by immunofluorescence whether soluble R28$_{Nt}$ and highlighted integrin monomers are in close proximity at the surface of HEC-1-A cells (Fig. 5A-5C). At the cell surface, R28$_{Nt}$ indeed binds into clusters in close proximity to integrin α3, α6, and β1 clusters.

To confirm a direct interaction between R28$_{Nt}$ and integrins, we analyzed the binding of the extracellular domains of integrins to R28$_{Nt}$, R28-N1 and R28-N2 by ELISA (Fig. 5D-5F). Integrins α3β1, α6β1 and α6β4 all interact significantly more with R28$_{Nt}$, R28-N1 and R28-N2 than with BSA (p<0.001). To assess the binding specificity further, the same experiment was carried out with an integrin containing a subunit which, although it is expressed by HEC-

1-A cells, had not been found in the mass spectroscopy analysis, namely α2β1 (Fig. S3) (32). In contrast to a natural ligand, type I collagen, neither R28$_{Nt}$ nor its two subdomains, R28-N1 and R28-N2, displayed interaction with α2β1, indicating that the integrins α3β1, α6β1 and α6β4 are true receptors of R28$_{Nt}$. The influence of divalent cations on the binding efficiency was assessed. As expected, binding of R28$_{Nt}$ to the different integrins in the presence of EDTA reflects that in PBS (Fig. 5G). Addition of the divalent ions Ca^{2+} significantly lowers binding of R28$_{Nt}$ to α3β1 and α6β1 (51% and 72% compared to in the presence of EDTA, respectively) and that of Mn^{2+} to α3β1 (45%). In contrast, binding to α6β4 was unchanged by the addition of divalent cations. Altogether, these data demonstrate the direct interaction between R28$_{Nt}$, R28-N1 and R28-N2 and the three integrins α3β1, α6β1 and α6β4.

R28$_{Nt}$ is sufficient to promote adhesion to different cell lines

So far, our study focused on the capacity of R28 to promote adhesion to endometritis-related cells; yet, *emm28* is the first to third most prevalent genotype for GAS invasive infections in Europe (9, 10, https://cnr-strep.fr/). Moreover, the R28-interacting integrins are expressed by numerous cell types, so we assayed whether R28 contributes to adhesion in other GAS-elicited invasive infections. We tested R28$_{Nt}$ ability to promote adhesion to other cell lines by testing the binding capacity of R28$_{Nt}$-coated beads to skin keratinocytes, HaCaT, pulmonary epithelial cell, A549, as well as the gut intestinal cells TC7, that do not display the integrins α3β1, α6β1 and α6β4 apically (Fig. 6) (35). R28$_{Nt}$-coated beads bind significantly and to a similar level (approximately 30%) A549, HaCaT and HEC-1-A cells (Fig. 2G and 6). In contrast, only 13% of R28$_{Nt}$-coated beads bind TC7 cells, not significantly different from the BSA-coated control beads. In conclusion, R28$_{Nt}$ is sufficient to promote adhesion to cell lines relevant not only for GAS endometritis but also for other GAS invasive infections.

Discussion

R28, a surface protein specifically expressed by *emm28* strains, which are associated with endometritis, promotes adhesion to cervical cells; its eukaryotic ligand was unknown (18). Currently in France, 50 % of GAS elicited endometritis are associated with puerperal fever, abortion, *in vitro* fertilization or spontaneous abortion (our unpublished data, https://cnr-strep.fr/), all situations where stromal decidual cells line the endometrium. Thus, GAS elicited endometritis may be favored by GAS adhering to decidual cells, in addition to epithelial cervical or endometrial cells. The R28 deleted strain was hampered in its capacity to adhere to human decidual stromal cells (hDSCs) (Fig. 2). That the mutant strain still adheres can be explained by the number of adhesins displayed by GAS strains (36–38), the deletion of a single one being insufficient for a total loss of adhesion capacity, as shown for example for Epf (39) and C5a peptidase (40). R28$_{Nt}$ binds to hDSCs, endometrial epithelial cells and cervical epithelial cells; it also promotes the adhesion of beads or heterologous bacteria to endometrial epithelial cells (Fig. 2). These data altogether confirm the role of R28 as an adhesin and extends its role to other cells potentially involved during endometritis.

R28 belongs to the Alp family of proteins, shared with GBS, which includes ACP, Rib and Alp2. The ACP N-terminal domain, but not that of Rib, also bound endometrial and cervical epithelial cells indicating that this property is not limited to R28 but not shared by all the members of the Alp family either (Fig. 3). Noteworthy, the ACP N-terminal domain shows a higher sequence identity to R28 N-terminal sequence than does Rib N-terminal domain, as well as a structural identity (18, 24); this could account for ACP but not Rib sharing the binding property. Furthermore, R28$_{Nt}$ and Alp2 N-terminal domain are near identical (20, 21), consequently this binding property could be extended to Alp2. Fifteen to 30% of women vaginal tract is colonized by GBS and all GBS strains present the gene encoding for one member of the Alp family (15, 41); most of these proteins are likely involved in GBS colonization.

The GBS proteins ACP and Alp3 (*i.e.* R28) bind glycosaminoglycan (26). Analysis using a set of constructions with different ACP domains suggested that a site beginning in the N-terminal region of ACP and extending into the repeat region is responsible for this binding (26, 27). Here, the R28$_{Nt}$ and ACP$_{Nt}$ peptides we used were devoid of repeats region sequences, suggesting that the binding we observed was not that to glycosaminoglycan; this was confirmed by the absence of heparinase treatment effect on the R28$_{Nt}$ − HEC-1-A binding. This suggested that

other receptors than glycosaminoglycan are involved in the R28-endometrial or -cervical cell binding.

In this study, pull-down experiments on whole cells indicate that R28$_{Nt}$ binds other integrins namely α3β1, α6β1 and α6β4 (Fig. 5). Furthermore, we demonstrate by protein-protein binding experiments that the binding is direct. Among already described GAS direct interactions, Scl1 interacts with integrins α11β1 and α2β1 (42, 43), SDH with urokinase plasminogen activator receptor (44), M6 and M1 with CD46 (45, 46) and hyaluronic acid capsule with CD44 (47). The direct binding of GAS to α3β1, α6β1 and α6β4 had never been described thus far. Cervical cells, endometrial epithelial cells and decidual cells all express, among others, the integrin subunits α3, α6, β1 and β4 (32, 48–50).The R28$_{Nt}$ interaction with integrins α3β1, α6β1 and α6β4 could therefore account for the binding of R28$_{Nt}$ to all these endometritis-related cells (Fig 2).

The α3β1, α6β1, and α6β4 may not be the sole R28$_{Nt}$ ligands. Indeed, we did not assay the direct interaction of R28$_{Nt}$ with all pulled-down proteins and the other candidates may interact indirectly, *via* the pulled-down integrins, or directly with R28$_{Nt}$. Other cell surface proteins could also interact with R28$_{Nt}$ and be missed in our screen since our co-immunoprecipitation experiments were carried out with HEC-1A cells; different cell types express different proteins, including different integrins. Further analyses may thus highlight other eukaryotic R28 receptors. Similarly, ACP binds α1β1 but may also bind other integrins (28) and among them α3β1, α6β1 or α6β4.

Our data suggesting that both R28-N1 and R28-N2 shared a receptor, we tested whether they both interacted directly with the three integrins, which they did (Fig. 5), supporting our competitive assay conclusion. To further characterize the interactions between R28$_{Nt}$ and integrins we assayed the influence of Mn^{2+} and Ca^{2+} on them. Our data suggested that at the molecular level the interactions between R28$_{Nt}$ and the various integrins are different and that cations induced conformational changes to α3β1 and α6β1, but not of α6β4, which interferes with R28$_{Nt}$ binding. Most integrin ligands require cations for their interactions, but, although not often described, cation independent interactions of integrins with proteins from pathogens have already been reported, for example between the

αVβ5 integrin and the HIV Tat protein (51). Defining the R28 amino-acid residues involved in the interaction would provide clues regarding the molecular mechanism taking place. Establishing the structure of R28-N2 would shed light on the sequences potentially interacting with the various receptors and provide a framework for mutational analysis. ACP binds α1β1 through a KTD sequence located in a β-sandwich subdomain (24, 28); however, R28 harbors no KTD or related sequence and interacts with other integrins, it will be interesting to determine whether the same β-sandwich subdomain is nevertheless involved (Fig. 1A).

The absence of significant competition between ACP and R28-N2 could be related to the strong difference of affinity between R28-N2 and ACP for HEC-1-A binding. Also, the absence of competition between R28-N1 and ACP-bound beads could be a consequence of ACP and R28-N1 presenting distinct affinity for some receptors or ACP recognizing additional receptors.

R28 promotes direct binding to the cells. However, this is not the sole binding involved in the adhesion of GAS with hDSC as indicated by the fact that the mutant strain still adheres to the cells. GAS adhesion is a two-step process, a first more labile adhesion involving lipoteichoic acid followed by a second more specific and stronger adhesion involving adhesins (52). Many adhesins have been described and the interactions with epithelial cell surface receptors thoroughly studied (reviewed 32, 47, 48). Most interactions are indirect, proteins of the extracellular matrix, such as fibronectin, form bridges with the host cells. GAS possesses numerous fibronectin-binding proteins, including the M protein that is variable among the *emm*-types and may not play always the exact same role. Also, M1 binds directly CD46 (46) and indirectly α5β1, *via* fibronectin. The dual interaction is required for efficient epithelial cell invasion (46). The interaction between R28 and integrins could potentiate other association(s) mediated by independent adhesins. More experimental support is needed to draw conclusions.

The involvement of R28 in the association between *emm28* and endometritis has been suggested (18) and our data support this suggestion. However, other factors may be involved in this association. They may be adhesins restricted to few genotypes, including *emm28*, such as Mrp28, M28, Enn28, FCT type 4 pilus proteins (Sfb1, Cpa) (53), Epf (39), and some RD2, a *emm28*-specific genomic region,

encoded surface proteins (13). The unique conjunction of these in *emm28* strains may confer an advantage for binding to the endometrium. RD2 also encodes other proteins of unknown function as well as a regulator and they may be involved in diverse colonization or invasion steps of the endometrial or cervical tissue. The adhesion, colonization and invasion capacities of single or multiple mutant strains should be assessed.

R28 promotes adhesion to female genital tract cells including epithelial cells and it was suggested that it may contribute to GAS adhesion to different types of epithelial cells (18). We have shown that, indeed, it also promotes adhesion to endometrial epithelial cells and decidual cells, but also to pulmonary epithelial cells and to keratinocytes, natural targets for GAS invasive infections (Fig. 6). HaCaT and A549 both express α3, α6, β4 and β1 at their surface (54, 55) and that R28$_{Nt}$ interacts with the different corresponding dimers probably accounts for the increase of binding of R28$_{Nt}$ beads to these cell lines (Fig. 6). TC7 cells also express these integrins, but their localization is mainly basolateral (35); this could explain the lower binding of R28$_{Nt}$ to the apical side of TC7 cells compared to the other cell lines. This ability of R28$_{Nt}$ to promote adhesion to several cell lines, at least partially through integrins, suggests that expressing R28 is an advantage for *emm28* strains that is not limited to endometritis. This may represent a physiological explanation for the fact that *emm28* genotype is a prevalent genotype, being the first to third most encountered *emm*-type in invasive infections, in most European countries (9, 10, https://cnr-strep.fr/). Other *emm28* specific virulence factors could also participate in the prevalence and warrant further study.

In conclusion, through direct interaction with integrins, R28 supports *emm28* strains elicited endometritis and also contributes to the prevalence of this genotype.

Experimental procedure:
Bacterial strains and growth conditions
The strains used in this study are described in Table S1. The M28PF1 strain is a clinical isolate responsible for an endometritis (French National Reference Center (CNR) for Streptococci, https://cnr-strep.fr/) that was selected on phenotypic and genotypic bases from a collection of 50 *emm28* independent clinical isolates (56). GAS strains were grown under static condition at 37°C in Todd Hewitt medium

supplemented with 0.2% Yeast Extract (THY) or on THY agar (THYA) plates. *L. lactis* strains were grown in Todd Hewitt (TH) medium, supplemented with 10 µg/mL erythromycin when necessary, at 30 °C without agitation or on TH agar (THA) plates. *Escherichia coli* strains were cultured in Tryptic Soy medium at 37 °C with agitation with, when necessary, added antibiotics at the following concentrations: erythromycin 150 µg/ml, ticarcillin, 100 µg/ml.

Genetic constructions and generation of the ΔR28 mutant
The DNA fragments encoding the proteins without signal peptide of R28$_{Nt}$, R28-N1, R28-N2, ACP and Rib$_{Nt+2R}$ (nt 169-1272, nt 169-688, nt 688-1272, nt 154-681 and nt 171-1188, respectively) were amplified and cloned into pET2818 as described in (57). All fragments were amplified by PCR using genomic DNA from M28PF1, except for the fragments encoding ACP and Rib$_{Nt+2R}$, which were amplified from GBS strains A909 and BM110, respectively; the primers used are described in Table S2. The plasmid pOri_R28$_{Nt}$, contains the promoter of *hvgA* deleted of the *covR* consensus boxes, the signal peptide of the GBS BM110 surface protein HvgA, and the LPXTG anchor signal of HvgA subcloned from pAT28-covSP +SPA (57), with the addition of the sequence encoding R28$_{Nt}$ subcloned from pET2818_R28$_{Nt}$. The different plasmids constructed and used during this study are described in Table S1.

The ΔR28 strain corresponds to an in-frame deletion mutant of the gene *Spy1336* encoding the R28 protein and was obtained by homologous recombination of the plasmid pG1-R28 following the same protocol as previously (57). The strain was entirely sequenced as described previously (56). No other significant mutation was found compared to the parent strain M28PF1 except for the gene deletion (56). The primers used for the generation of the plasmid pG1-R28 are described in Table S2. *L. lactis* strain was transformed by electroporation (58) with the empty vector (CCH2022) or pOri_R28$_{Nt}$ (CCH2023).

Protein production and purification
Peptide expressions were performed with the corresponding derived pET2818 plasmids (Table S1) and purification as described in (57), except for the gel filtration step which we did not perform. Peptides purity were confirmed using a

12% SDS-PAGE acrylamide gel and Coomassie Blue staining (Fig. S1C).

Antibodies

All animal experiments described in this study were conducted in accordance with guidelines of Paris Descartes University, in compliance with the European animal welfare regulation (http://ec.europa.eu/environment/chemicals/lab_animals/home_en.html) and were approved by the Institut Cochin University Paris Descartes animal care and use committee (n° 12–145). Mouse anti-$R28_{Nt}$ anti-serum was produced as follows. BalbC 6 weeks mice (Janvier laboratory) were subcutaneously injected with 100 µL of 50 µg/mL $R28_{Nt}$ solution in PBS+aluminium adjuvant. Two and four weeks after initial injection, mice were injected with 100 µL of a 25 µg/mL solution. The antibody purity and titer were determined a week after the last injection and assessed by western blot experiments. The purified R28-N1 or R28-N2 were injected into rabbits to produce polyclonal antibodies (Covalab). The following antibodies were used throughout the study: anti α3 integrin (P1B5; DSHB), rabbit anti α6 integrin (Novus) for ELISA, mouse anti α6 integrin (P5G10; DSHB) for immunofluorescence, mouse anti β1 integrin (P5D2; DSHB) and rat anti β1 integrin (AIIB2, DSHB).

Bacterial cell wall extracts

Overnight cultures of *L. lactis* were diluted 1/100 in 50 mL of TH broth and cultivated to DO = 0.5. Bacteria were centrifuged 10 min at 5000 rpm at 4°C and the pellet was washed once in wash buffer (Cold PBS 1X + EDTA 10 mM + PMSF 1 mM). Bacteria were centrifuged and the pellet was resuspended in 250 µL of mutanolysin buffer (Tri-HCL 20 mM pH 7.5 + Sucrose 1 M + EDTA 10 mM + PMSF 1 mM + protease inhibitor cocktail 1X (Roche) + mutanolysin 200 u/mL (Sigma)) and incubated for 90 minutes at 37°C. The protoplast suspension is then centrifuged 15 minutes at 10 000 rpm and 30 µL of the supernatant containing the cell wall extract is loaded on an acrylamide gel.

Cell culture

HEC-1-A (ATCC® HTB-112™) and ME-180 (ATCC® HTB-33™) cells were cultured as recommended, in McCoy's 5A medium (GibCo) supplemented with 10% fetal bovine serum (FBS) at 37 °C, 5% CO_2. A549 (ATCC® CCL-185™), HaCaT (AddexBio T0020001) and Caco/TC7 cells (CVCL_0233), were cultivated as recommended, in RPMI + 10 % FBS.

Isolation and culture of human decidual stromal cells (hDSC)

Decidual stromal cells were isolated from decidua parietalis, obtained from fetal membranes of non laboring women after a normal term (> 37 weeks of gestation) singleton-pregnancy delivered by elective caesarean section. The study of the human fetal membranes was approved by the local ethics committee (Comité de Protection des Personnes Ile de France XI, n° 11018, 03/03/2011) and informed consent was obtained from all donors. Furthermore, the study abides to the Declaration of Helsinki principles. Briefly, fetal membranes were dissected from placenta under sterile conditions and decidua attached to chorion leaf was peeled off amnion and placed in PBS. After the removal of blood clots, choriodecidua was cut in small pieces and digested with 0.2% collagenase B (Roche Diagnostics, Mannheim, Germany) in DMEM-F12 (Invitrogen, Cergy-Pontoise, France) at 37°C for 1h30. After the addition of DMEM-F12 containing 5% FCS and 100 µM EDTA, the cell suspension was filtered through a 100 µm nylon gauze and centrifuged at 400 g for 10 min. The cell pellet, resuspended in complete medium (DMEM-F12 containing 5% FCS, 100 IU/ml penicillin (Invitrogen), 100 µg/ml streptomycin (Invitrogen)), were plated at a density of 10^5 cells/cm² and cultured in complete medium at 37°C in 5% CO_2 and 95% air in 75cm² flasks overnight. Medium was then renewed after several PBS washes. Cells were harvested with trypsin/EDTA when the cells were 90% confluent. The hDSCs were expanded further to passage 2 to 4 and thereafter, used in the experiments. hDSCs from at least two different women were used.

Adhesion of bacteria to cells

GAS strains M28PF1 (WT) and ΔR28 were diluted from an overnight culture, grown to the exponential phase at an OD = 0.5 and diluted in RPMI without glutamine to obtain a multiplicity of infection (MOI) of one. Confluent hDSCs in 24 well-plates were washed three times, bacterial solution was added and plates were centrifuged 5 min at 1000 rpm to synchronize bacterial adhesion. After 1 h of incubation at 4 °C, cells were washed three times with PBS and lysed with distilled water. Serial dilutions of cellular lysates were plated on THYA plates and the number of CFUs was determined after 24-48

hours growth at 37 ℃. Six independent experiments were performed in triplicates and ΔR28 values were normalized to WT adhesion for each independent experiment. Statistical analysis was performed using One-Sample T-test with Graph-Pad Prism 5.0.

For adhesion assay of *Lactococcus lactis* on HEC-1-A, cells were seeded 72 h prior to infection to reach confluence at 5-10 x 10^6 cells/well in 24-well plates. The protocols for the *L. lactis* CCH2022 and CCH2023 bacterial preparation and adhesion to cells was the same as above. For all experiments, three independent assays in triplicate were carried out for each infection. Paired T-test were performed with Graph-Pad Prism 5.0.

ELISA-based protein-cell interaction assay

The cell-protein binding assay protocol derives from Bolduc et *al.* 2007 (28). Briefly, cells are incubated with biotinylated proteins in PBS + BSA 1% for 1 hour at 4°C, then washed and bound proteins are revealed by streptavidin-HRP (GE Healthcare) and subsequent O-phenyldiamine (Sigma) revelation. Number of moles of biotin per mole of protein was evaluated using an HABA kit (Lifetechnologies): the ratio of biotin molecule per peptide molecule was around 5, except for Rib_{Nt+2R} for which it reached 9. Therefore, the fluorescence values obtained for Rib_{Nt+2R} were divided by 2 for comparison sake. Experiments were performed four times independently, except for Rib_{Nt+2R} performed three times. Statistical analysis was performed after non-linear fitting using a second order polynomial quadratic model and comparison of the extra sum of squares F-test of the best fit values using GraphPad Prism 5.0. Curves are fitted with an equation: Absorbance = B1 * concentration+ B2* (concentration)². Since this equation did not fit the BSA curve, statistical analysis to BSA were performed with Two-Way ANOVA with Bonferroni post-tests at maximum binding.

Beads fixation to cells

Fluosphere (Invitrogen, yellow green, 1 µm) were coupled with the different peptides following the manufacturer's recommendations. Correct coupling was confirmed by cytometric analysis with appropriate antisera. Indicated confluent cells in dark 96 well plates (Nunc, Lifetechnologies) were washed three times with cold PBS supplemented with 1 mM Ca^{2+} and 1 mM Mg^{2+} (PBS++). Beads were diluted to a concentration of 10^8 beads/mL in cold PBS++ and 150 µL of beads solution was added to cells and incubated 1 h at 4 °C. Cells were washed three times in PBS++, fixed 15 min at room temperature in paraformaldehyde 1%, quenched with 50 mM NH_4Cl. Fluorescence of the inoculum solutions, cells and cells incubated with beads were measured with a TECAN fluorescent plate reader with a GFP filter. Fluorescence of adherent beads corresponds to the fluorescence of the total well subtracted by the autofluoresence of a well containing cells only.

For competition assays, prior to beads adhesion, cells were incubated with 20 µM of specified peptide in PBS++ for 1 h at 4 °C; beads were then added and the experiments performed as described above. For antibody competition, beads were incubated with 10 µg/mL of purified anti-R28-N1 and anti-R28-N2 rabbit antibodies for 1 h at 4°C prior to incubation with cells. Experiments were performed at least three times in duplicates. Statistical analyses were performed with Two-Way ANOVA.

Flow cytometry analysis of protein binding to eukaryotic cells after different treatments

Confluent 72 h HEC-1-A cells were dissociated with cell dissociation buffer (Lifetechnologies) and treated with $NaIO_4$ (Sigma), pronase (Lifetechnologies), trypsin (Worthington Biochemical Corporation), phospholipase A_2 (Sigma) or heparinase I (NEB) at the specified concentrations and as described by Gallotta et *al.* 2014 (33). Cells were then incubated with 20 µM soluble $R28_{Nt}$, R28-N1 or R28-N2 in PBS- BSA 0,5% and stained with mouse antiserum, and secondary anti-mouse PE coupled antibody. Fluorescence of cells was measured in an Accuri C6 BD cytometer. 100% staining was considered with cells untreated (Control). Unspecific labelling of cells by mouse antiserum and secondary antibodies was measured on untreated cells that were not incubated with the peptides, but underwent the same labeling protocol (without peptide condition). Experiments were performed at least three times independently. Two-Way ANOVA analysis was performed with GraphPad Prism 5.

Protein sequence comparison

These comparisons were carried out using the BlastP algorithm (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE= Proteins&PROGRAM=blastp&BLAST_PROGR AMS=blastp&PAGE_TYPE=BlastSearch&BLA

ST_SPEC=blast2seq&DATABASE=n/a&QUERY=&SUBJECTS=)

In silico *protein structure prediction*

Peptide structure was made using the Phyre software (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index).

Co-immunoprecipitation of cross-linked R28_{Nt} receptors

HEC-1-A cells grown to confluence in T25 flask were washed three times with cold PBS++. A 6 mL solution of cold PBS++, 1% BSA, with or without (control) 20 µM R28_{Nt} was added and incubation pursued for 1 h on ice. Cells were washed three times in cold PBS++ and 10 mL of 1 mM DTSSP (Thermofisher) was added for 2 h on ice. After one wash with PBS++, cross-linking was blocked with 10 mL of 50 mM Tris pH 7.5 for 10 min on ice. An equimolar mix of rabbit polyclonal anti R28-N1 and R28-N2 antibodies (1 mg/mL total, 1.7 mL per flask) was added to two freshly washed HEC-1-A T25 flask for 10 min to adsorb unspecific cell binding antibodies. Pre-adsorbed antibodies (1.5 mL) were added to each flask and incubated for 1 h at 4 °C. Unbound antibodies were removed by three washes with PBS++ and cells were lysed for 1 h at 4 °C with 1.5 mL of lysis buffer (PBS, protease inhibitor cocktail 1X (Thermofisher), phenylmethylsulfonyl fluoride 1 mM (Sigma), 1% SurfactAmp Triton X100 (Thermofisher)). Lysates were centrifuged at 15 000 g for 15 min at 4 °C and supernatants were incubated with 100 µL magnetic protein A beads (Pierce) for 1 h at 20 °C. Beads were then washed six times with 1.5 mL PBS, 500 mM NaCl, Triton 1% and complexes were eluted in Laemmli Buffer in disulfide-bonds preservation conditions (60mM Tris pH6.8, 1% SDS) at 95°C for 5 minutes. Gels were run in duplicate; one gel was used for western blotting with 15 % of the eluates, revealed with mouse anti-R28_{Nt} antiserum to reveal cross-linked complexes and the other one for their identification by mass spectrometry. The western blotting indicated that R28_{Nt} complexes were found above 170 kDa.

Short migration SDS-PAGE, protein trapping and peptide extraction

Eighty-five % of the eluate were loaded in the same condition as above for a two centimeters-migration, then stained with colloidal Coomassie blue (Quick Coomassie Stain from Cliniscences).

Stained gels allowed visualization of protein abundance and molecular weight distribution. Protein-containing gel lanes from the two conditions (with the bait R28 and in its absence, control condition) were excised above the 170 kDa MW marker (Lifetechnologies Pageruler prestained protein ladder). The polyacrylamide gel constitutes a matrix where successive protein treatments were performed: salt, buffer and detergent removal by successive washes of 100 mM NH_4HCO_3 (or ABC)/Acetonitrile 50%, disulfide bonds removal by cystein reduction in ABC + 10 mM DTT at 56 °C for 30 minutes; free thiols protection by alkylation in ABC + 55 mM chloroacetamide for 30 minutes at room temperature and overnight digestion with trypsin. Peptides were then extracted using washes of 5% formic acid intercalated with gel shrinkages in 50% acetonitrile (ACN). All washes were pooled and evaporated before analysis.

C18 liquid nanochromatography and mass spectrometry

Mass spectrometry analyses were performed at the 3P5 proteomics facility of the University Paris Descartes using an U3000 RSLC nano-LC-system hyphenated to an Orbitrap-fusion mass spectrometer, all from Thermo Fisher Scientific. Peptides were solubilized in 7 µl of 0.1% TFA containing 10% ACN. They were loaded, concentrated and washed for 3 min on a C18 reverse-phase precolumn (3 µm particle size, 100 Å pore size, 75 µm inner diameter, 2 cm length, Thermo Fischer Scientific). Peptides were separated on a C18 reverse-phase resin (2 µm particle size, 100 Å pore size, 75 µm inner diameter, 25 cm length from Thermofisher scientific) with a 35 minutes binary gradient starting from 99% of solvent A containing 0.1% formic acid in H_2O and ending in 40% of solvent B containing 80% acetonitrile, 0.085% formic acid in H_2O. The mass spectrometer acquired data throughout the elution process and operated in a data-dependent scheme with full MS scans acquired with the Orbitrap, followed by as many MS/MS ion trap CID spectra 5 seconds can fit (data-dependent acquisition with top speed mode: 5 seconds cycle) using the following settings: full MS automatic gain control (AGC): $2.10e^5$, maximum ion injection time (MIIT): 60 ms, resolution: $6.10e^4$, m/z range 350-1500 and for MS/MS; isolation width: 1.6 Th, minimum signal threshold: 5000, AGC $2.10e^4$, MIIT: 100 ms, peptides with undefined charge state or charge

state of 1 were excluded from fragmentation, dynamic exclusion time: 30 s.

Identifications (protein hits) and quantifications were performed by comparison of experimental peak lists with a database of theoretical sequences using MaxQuant version 1.6.1.0. (59). The databases used were the human sequences from the curated Uniprot database (release june 2018) and a list of in-house contaminant sequences. Carbamidomethylation of cysteins was set as constant modification while acetylation of protein N-terminus and oxidation of methionines were set as variable modifications. False discovery rate (FDR) was kept below 1% on both peptides and proteins. The "match between runs" option was not allowed. For analysis, results from MaxQuant were imported into the Perseus software version 1.6.1.1. (60). Reverse and contaminant proteins (human keratins and non-human proteins) were excluded from analysis. Proteins of interest were selected based on reported intensity in 3 over 3 replicates for each group. All hits are presented in the supplementary file.

Selection of positive hits

The enrichment is defined as the ratio of raw intensity of the co-immunoprecipitation identified proteins in the R28_{Nt} bait conditions compared to the control condition. Hits with a ratio above 1 and/or no peptide found in the control condition for all three independent experiments performed were considered as specifically enriched in the presence of R28. From this list, only surface exposed proteins are shown in Table 1.

Immunolabelling of bound R28_{Nt} and coated fluospheres on HEC-1-A

Seventy-two hours after seeding on glass coverslips in 24 well plates, confluent HEC-1-A cells were washed three times in cold PBS++. A 10 µM solution of R28_{Nt} in PBS++, 1% BSA was added to the cells for 1 h at 4 °C. Cells were washed three times in cold PBS++ and fixed 15 min at 20 °C with paraformaldehyde 1%. Cells were quenched with ammonium chloride 50 mM and blocked for 30 min with PBS-BSA 3%. Cells were incubated with primary antibodies for 1 h at 20 °C in PBS-BSA 1%. Cells were washed and incubated with secondary antibodies and DAPI at 1/5 000 for 1 h at 20 °C. Coverslips were mounted on slides with Mowiol, imaged on a Leica DMI6000 and images were analyzed with ImageJ. For the observation of bound coated fluospheres, 10^8 coated beads in PBS++ were added to the cells and treated as above. Beads did not require labelling since they are intrinsically fluorescent.

ELISA

To evaluate integrin binding to immobilized R28_{Nt} or its sub-domains R28-N1 and R28-N2, ELISA were performed essentially as described in Six et *al.* 2015 (57). Proteins were coated at 5 µg/mL overnight. Fifty µL of integrins (human recombinant extracellular domain, tested for their ability to interact with known ligands, R&D) were incubated 2 h at 37 °C at the specified concentrations diluted in PBS-BSA 1%. Rat anti-β1 (AIIB2, DSHB) and rabbit anti-α6 (R&D) antibodies were used at a 1/200 dilution. Experiments were performed at least three times. Statistical analyses were performed as described above for the ELISA based protein-cell interaction assay.

The importance of divalent cations for the interaction between R28_{Nt} and integrins was tested by ELISA with the same coating protocol as above with fifty µL at 10 µg/mL of integrins added to the wells. All solutions for the assay were made in PBS, PBS + 10 mM EDTA, PBS + 1 mM Mn^{2+} and PBS + 1 mM Ca^{2+}, as specified. Detection and revelation were performed as above. Values were normalized on the binding to the integrins in the presence of EDTA. One-Way ANOVA was performed to compare the four independent experiments.

Conflict of interest: The authors declare none

References

1. Bisno, A. L., Brito, M. O., and Collins, C. M. (2003) Molecular basis of group A streptococcal virulence. Lancet Infect. Dis. 3, 191–200
2. Tart, A. H., Walker, M. J., and Musser, J. M. (2007) New understanding of the group A Streptococcus pathogenesis cycle. Trends Microbiol. 15, 318–25
3. Cunningham, M. W. (2000) Pathogenesis of group A streptococcal infections. Clin Microbiol Rev. 13, 470–511
4. Carapetis, J. R., Steer, A. C., Mulholland, E. K., and Weber, M. (2005) The global burden of group A streptococcal diseases. Lancet Infect. Dis. 5, 685–94
5. Beall, B., Facklam, R., and Thompson, T. (1995) Sequencing emm -specific polymerase chain reaction products for routine and accurate typing of group A Streptococci. J.Clin.Microbiol. 34, 953–958
6. Ferretti, J. J., Stevens, D. L., and Fischetti, V. A. (2016) Streptococcus pyogenes Basic Biology to Clinical Manifestations
7. Bessen, D. E. (2016) Tissue tropisms in group A Streptococcus: what virulence factors distinguish pharyngitis from impetigo strains? Curr. Opin. Infect. Dis. 29, 295–303
8. Beres, S. B., and Musser, J. M. (2007) Contribution of exogenous genetic elements to the group A Streptococcus metagenome. PLoS One. 10.1371/journal.pone.0000800
9. Gherardi, G., Vitali, L. A., and Creti, R. (2018) Prevalent emm Types among Invasive GAS in Europe and North America since Year 2000. Front. Public Heal. 6, 59
10. Plainvert, C., Doloy, A., Loubinoux, J., Lepoutre, A., Collobert, G., Touak, G., Trieu-Cuot, P., Bouvet, A., and Poyart, C. (2012) Invasive group A streptococcal infections in adults, France (2006-2010). Clin. Microbiol. Infect. 18, 702–10
11. Green, N. M., Beres, S. B., Graviss, E. A., Allison, J. E., McGeer, A. J., Vuopio-Varkila, J., LeFebvre, R. B., and Musser, J. M. (2005) Genetic diversity among type emm28 group A Streptococcus strains causing invasive infections and pharyngitis. J. Clin. Microbiol. 43, 4083–91
12. Colman, G., Tanna, A., Efstratiou, A., and Gaworzewska, E. T. (1993) The serotypes of Streptococcus pyogenes present in Britain during 1980-1990 and their association with disease. J. Med. Microbiol. 39, 165–178
13. Green, N. M., Zhang, S., Porcella, S. F., Nagiec, M. J., Barbian, K. D., Beres, S. B., LeFebvre, R. B., and Musser, J. M. (2005) Genome sequence of a serotype M28 strain of group a streptococcus: potential new insights into puerperal sepsis and bacterial disease specificity. J. Infect. Dis. 192, 760–70
14. Sitkiewicz, I., Green, N. M., Guo, N., Mereghetti, L., and Musser, J. M. (2011) Lateral gene transfer of streptococcal ICE element RD2 (region of difference 2) encoding secreted proteins. BMC Microbiol. 11, 65
15. Baker, C. J., Goroff, D. K., Alpert, S., Crockett, V. A., Zinner, S. H., Evrard, J. R., Rosner, B., and McCormack, W. M. (1977) Vaginal colonization with group B streptococcus: a study in college women. J. Infect. Dis. 135, 392–7
16. Nobbs, A. H., Lamont, R. J., and Jenkinson, H. F. (2009) Streptococcus adherence and colonization. Microbiol. Mol. Biol. Rev. 73, 407–50, Table of Contents
17. Courtney, H. S., Hasty, D. L., and Dale, J. B. (2002) Molecular mechanisms of adhesion, colonization, and invasion of group A streptococci. Ann. Med. 34, 77–87
18. Stålhammar-Carlemalm, M., Areschoug, T., Larsson, C., and Lindahl, G. (1999) The R28 protein of Streptococcus pyogenes is related to several group B streptococcal surface proteins, confers protective immunity and promotes binding to human epithelial cells. Mol. Microbiol. 33, 208–19
19. Wästfelt, M., Stålhammar-Carlemalm, M., Delisse, A. M., Cabezon, T., and Lindahl, G. (1996) Identification of a family of Streptococcal surface proteins with extremely repetitive structure. J. Biol. Chem. 271, 18892–18897
20. Lachenauer, C. S., Creti, R., Michel, J. L., and Madoff, L. C. (2000) Mosaicism in the alpha-like protein genes of group B streptococci. Proc. Natl. Acad. Sci. U. S. A. 97, 9630–9635
21. Lindahl, G., Stalhammer-Carlemalm, M., and Areschoug, T. (2005) Surface Proteins of Streptococcus agalactiae and Related Proteins in Other Bacterial Pathogens. Clin. Microbiol. Rev. 18, 102–127
22. Callebaut, I., Gilgès, D., Vigon, I., and Mornon, J. P. (2000) HYR, an extracellular module involved in cellular adhesion and related to the immunoglobulin-like fold. Protein Sci. 9, 1382–90
23. Glaser, P., Rusniok, C., Buchrieser, C., Chevalier, F., Frangeul, L., Msadek, T., Zouine, M., Couvé, E.,

Lalioui, L., Poyart, C., Trieu-Cuot, P., and Kunst, F. (2002) Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease. Mol. Microbiol. 45, 1499–1513

24. Aupérin, T. C., Bolduc, G. R., Baron, M. J., Heroux, A., Filman, D. J., Madoff, L. C., and Hogle, J. M. (2005) Crystal structure of the N-terminal domain of the group B streptococcus alpha C protein. J. Biol. Chem. 280, 18245–52

25. Stålhammar-Carlemalm, M., Areschoug, T., Larsson, C., and Lindahl, G. (2000) Cross-protection between group A and group B streptococci due to cross-reacting surface proteins. J. Infect. Dis. 182, 142–149

26. Baron, M. J., Bolduc, G. R., Goldberg, M. B., Aupérin, T. C., and Madoff, L. C. (2004) Alpha C protein of group B Streptococcus binds host cell surface glycosaminoglycan and enters cells by an actin-dependent mechanism. J. Biol. Chem. 279, 24714–24723

27. Baron, M. J., Filman, D. J., Prophete, G. A., Hogle, J. M., and Madoff, L. C. (2007) Identification of a glycosaminoglycan binding region of the alpha C protein that mediates entry of group B Streptococci into host cells. J. Biol. Chem. 282, 10526–10536

28. Bolduc, G. R., and Madoff, L. C. (2007) The group B streptococcal alpha C protein binds α1β1-integrin through a novel KTD motif that promotes internalization of GBS within human epithelial cells. Microbiology. 153, 4039–4049

29. Bolduc, G. R., Baron, M. J., Gravekamp, C., Lachenauer, C. S., and Madoff, L. C. (2002) The alpha C protein mediates internalization of group B Streptococcus within human cervical epithelial cells. Cell Microbiol. 4, 751–758

30. Chapot-Chartier, M. P., Vinogradov, E., Sadovskaya, I., Andre, G., Mistou, M. Y., Trieu-Cuot, P., Furlan, S., Bidnenko, E., Courtin, P., Péchoux, C., Hols, P., Dufre, Y. F., and Kulakauskas, S. (2010) Cell surface of Lactococcus lactis is covered by a protective polysaccharide pellicle. J. Biol. Chem. 285, 10464–10471

31. Areschoug, T., Carlsson, F., Stålhammar-Carlemalm, M., and Lindahl, G. (2004) Host-pathogen interactions in Streptococcus pyogenes infections, with special reference to puerperal fever and a comment on vaccine development. Vaccine. 22 Suppl 1, S9–S14

32. Castelbaum, A. J., Ying, L. E. I., Somkuti, S. G., Sun, J., Ilesanmi, A. O., and Lessey, B. a (1997) Characterization of Integrin Expression in a Well Differentiated Endometrial Adenocarcinoma Cell Line (Ishikawa). J. Clin. Endocrinol. Metab. 82, 136–142

33. Gallotta, M., Gancitano, G., Pietrocola, G., Mora, M., Pezzicoli, A., Tuscano, G., Chiarot, E., Nardi-Dei, V., Taddei, A. R., Rindi, S., Speziale, P., Soriani, M., Grandi, G., Margarit, I., and Bensi, G. (2014) SpyAD, a moonlighting protein of group A Streptococcus contributing to bacterial division and host cell adhesion. Infect. Immun. 82, 2890–2901

34. Luca-Harari, B., Darenberg, J., Neal, S., Siljander, T., Strakova, L., Tanna, A., Creti, R., Ekelund, K., Koliou, M., Tassios, P. T., and Van Der Linden, M. (2009) Clinical and Microbiological Characteristics of Severe Streptococcus pyogenes Disease in Europe. J. Clin. Microbiol. 47, 1155–1165

35. Hamzaoui, N., Kernéis, S., Caliot, E., and Pringault, E. (2004) Expression and distribution of [beta]1 integrins in in vitro-induced M cells: Implications for Yersinia adhesion to Peyer's patch epithelium. Cell. Microbiol. 6, 817–828

36. Nobbs, A. H., Jenkinson, H. F., and Everett, D. B. (2015) Generic determinants of Streptococcus colonization and infection. Infect. Genet. Evol. 33, 361–370

37. Brouwer, S., Barnett, T. C., Rivera-Hernandez, T., Rohde, M., and Walker, M. J. (2016) Streptococcus pyogenes adhesion and colonization. FEBS Lett. 10.1002/1873-3468.12254

38. Rohde, M., and Cleary, P. P. (2016) Adhesion and invasion of Streptococcus pyogenes into host cells and clinical relevance of intracellular streptococci. in Streptococcus pyogenes: Basic Biology to Clinical Manifestations, pp. 1–39

39. Linke, C., Siemens, N., Oehmcke, S., Radjainia, M., Law, R. H. P., Whisstock, J. C., Baker, E. N., and Kreikemeyer, B. (2012) The extracellular protein factor epf from Streptococcus pyogenes is a cell surface adhesin that binds to cells through an n-terminal domain containing a carbohydrate-binding module. J. Biol. Chem. 287, 38178–38189

40. Lynskey, N. N., Reglinski, M., Calay, D., Siggins, M. K., Mason, J. C., Botto, M., and Sriskandan, S. (2017) Multi-functional mechanisms of immune evasion by the streptococcal complement inhibitor C5a peptidase. PLoS Pathog. 10.1371/journal.ppat.1006493

41. Kong, F., Gowan, S., Martin, D., James, G., and Gilbert, G. L. (2002) Molecular Profiles of Group B

Streptococcal Surface Protein Antigen Genes: Relationship to Molecular Serotypes. J. Clin. Microbiol. 40, 620–626

42. Humtsoe, J. O., Kim, J. K., Xu, Y., Keene, D. R., Höök, M., Lukomski, S., and Wary, K. K. (2005) A streptococcal collagen-like protein interacts with the α2β1 integrin and induces intracellular signaling. J. Biol. Chem. 280, 13848–13857

43. Caswell, C. C., Barczyk, M., Keene, D. R., Lukomska, E., Gullberg, D. E., and Lukomski, S. (2008) Identification of the first prokaryotic collagen sequence motif that mediates binding to human collagen receptors, integrins α 2β1 and α11β1. J. Biol. Chem. 283, 36168–36175

44. Jin, H., Song, Y. P., Boel, G., Kochar, J., and Pancholi, V. (2005) Group A streptococcal surface GAPDH, SDH, recognizes uPAR/CD87 as its receptor on the human pharyngeal cell and mediates bacterial adherence to host cells. J. Mol. Biol. 350, 27–41

45. Okada, N., Liszewski, M. K., Atkinson, J. P., and Caparon, M. (1995) Membrane cofactor protein (CD46) is a keratinocyte receptor for the M protein of the group A streptococcus. Proc. Natl. Acad. Sci. U. S. A. 92, 2489–93

46. Rezcallah, M. S., Hodges, K., Gill, D. B., Atkinson, J. P., Wang, B., and Cleary, P. P. (2005) Engagement of CD46 and alpha5beta1 integrin by group A streptococci is required for efficient invasion of epithelial cells. Cell. Microbiol. 7, 645–53

47. Cywes, C., Stamenkovic, I., and Wessels, M. R. (2000) CD44 as a receptor for colonization of the pharynx by group A Streptococcus. J. Clin. Invest. 106, 995–1002

48. Werner, J., Decarlo, C. A., Escott, N., Zehbe, I., and Ulanova, M. (2012) Expression of integrins and Toll-like receptors in cervical cancer: Effect of infectious agents. Innate Immun. 18, 55–69

49. Lessey, B. a, Castelbaum, a J., Buck, C. a, Lei, Y., Yowell, C. W., and Sun, J. (1994) Further characterization of endometrial integrins during the menstrual cycle and in pregnancy. Fertil. Steril. 62, 497–506

50. Wadehra, M., Forbes, A., Pushkarna, N., Goodglick, L., Gordon, L. K., Williams, C. J., and Braun, J. (2005) Epithelial membrane protein-2 regulates surface expression of αvβ3 integrin in the endometrium. Dev. Biol. 287, 336–345

51. Vogel, B. E., Lee, S.-J., Hilderbrand, A., Craig, W., Pierschbacher, M. D., Wong-Staal, F., and Ruoslahti, E. (1993) A novel integrin specifically exemplified by binding avb5 integrin to the basic domain of the HIV tat protein anmd vitronectin. J.Cell Biol. 121, 461–468

52. Hasty, D. L., Ofek, I., Courtney, H. S., and Doyle, R. J. (1992) Multiple adhesins of streptococci. Infect. Immun. 60, 2147–2152

53. Köller, T., Manetti, A. G. O., Kreikemeyer, B., Lembke, C., Margarit, I., Grandi, G., and Podbielski, A. (2010) Typing of the pilus-protein-encoding FCT region and biofilm formation as novel parameters in epidemiological investigations of Streptococcus pyogenes isolates from various infection sites. J. Med. Microbiol. 59, 442–52

54. Abban, C. Y., and Meneses, P. I. (2010) Usage of heparan sulfate, integrins, and FAK in HPV16 infection. Virology. 403, 1–16

55. Falcioni, R., Cimino, L., Gentileschi, M. P., D'Agnano, I., Zupi, G., Kennel, S. J., and Sacchi, A. (1994) Expression of [beta]1, [beta]3, [beta]4, and [beta]5 Integrins by Human Lung Carcinoma Cells of Different Histotypes. Exp. Cell Res. 210, 113–122

56. Longo, M., De Jode, M., Plainvert, C., Weckel, A., Huaa, A., Château, A., Glaser, P., Poyart, C., and Fouet, A. (2015) Complete Genome Sequence of Streptococcus pyogenes emm28 strain M28PF1, responsible of a puerperal fever. Genome Announc.

57. Six, A., Bellais, S., Bouaboud, A., Fouet, A., Gabriel, C., Tazi, A., Dramsi, S., Trieu-Cuot, P., and Poyart, C. (2015) Srr2, a multifaceted adhesin expressed by ST-17 hypervirulent Group B Streptococcus involved in binding to both fibrinogen and plasminogen. Mol. Microbiol. 97, 1209–1222

58. Caparon, M. G., and Scott, J. R. (1991) Genetic Manipulation of Pathogenic Streptococci. Methods Enzymol. 204, 556–586

59. Cox, J., Hein, M. Y., Luber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014) Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Mol. Cell. Proteomics. 13, 2513–2526

60. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods. 13, 731–740

FOOTNOTES

The abbreviations used are: GAS, Group A Streptococcus; GBS, Group B Streptococcus.

Table 1. Main hits from mass spectrometry analysis of specifically co-immunoprecipitated R28$_{Nt}$ receptors

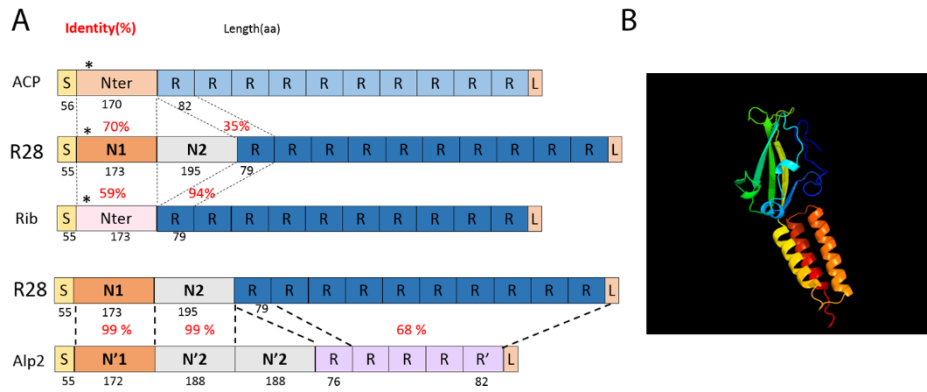| Common Name	Gene name	Relative raw intensity		
		Exp1	Exp2	Exp3
Basal cell adhesion molecule	*BCAM*	A	3.50	5.00
Cadherin-1;	*CDH1*	A	A	A
CD166 antigen	*ALCAM*	A	3.70	8.30
Desmoglein-2	*DSG2*	A	A	6.01
Ephrin type-A receptor 2	*EPHA2*	A	4.03	3.20
Integrin alpha-3	*ITGA3*	A	8.32	65.32
Integrin alpha-6	*ITGA6*	A	18.37	A
Integrin alpha-V	*ITGAV*	A	1.74	7.20
Integrin beta-1	*ITGB1*	A	3.50	6.55
Integrin beta-4	*ITGB4*	A	5.84	5.62
Prostaglandin F2 receptor negative regulator	*PTGFRN*	A	A	A
Transferrin receptor protein 1	*TFRC*	A	3.54	5.10

A, Absent from control condition.

Figure 1. **Schematic representation of R28, Alp family members and R28-N1 predicted 3D structure.** *A,* Schematic representation of the Alp family members structure and similarity. Numbers in dark below the scheme indicate the length in amino-acid residues and in red the percentage of identify between two domains. S=Signal peptide, R= repeats, L= LPXTG. *= KTD in ACP, KAD in R28 and KPD in Rib. Modified with our sequence comparisons from (1–3). *B,* R28-N1 structure as predicted by Phyre. Green and yellow, the β-sandwich; yellow to red, the three-helix bundle

Figure 2. R28$_{Nt}$ is sufficient to promote adhesion to female genital tract cells. *A*, Binding of the ΔR28 strain to human decidual stromal cells (DSCs) is expressed as a percentage of that of the wild-type strain; Two-tailed t-test was performed on 6 independent experiments performed in triplicates. *B-D,* ELISA-based protein-cell interaction assay with biotinylated proteins, expressed as percentage of the maximum binding, on: *B,* secondary decidual cells (DSCs); *C,* ME180 cells and *D,* HEC-1-A cells. Circle, biotinylated R28$_{Nt}$; square, biotinylated BSA. Error bars correspond to SEM of four independent experiments. *E-F* representative immunofluorescence of coated fluorescent beads with *E,* R28$_{Nt}$ or *F,* BSA*,* on HEC-1-A cells. Coated beads in green and DAPI staining in blue. *G,* Binding of R28$_{Nt}$ or BSA coated fluorescent beads to HEC-1-A cells. One-tailed t-test. *H,* Binding of R28$_{Nt}$-coated fluorescent beads to HEC-1-A cells in the presence of purified anti-R28$_{Nt}$ antibodies, expressed as the percentage of that in the absence of anti-R28$_{Nt}$ antibodies. *I,* Adhesion of R28$_{Nt}$ expressing *L. lactis* (L.l + R28$_{Nt}$) to HEC-1-A cells, expressed as the percentage of that of the *L. lactis* empty vector strain (L.l. + 0). *H-I,* two-tailed t-test. *, $p < 0.05$; **, $p < 0.01$
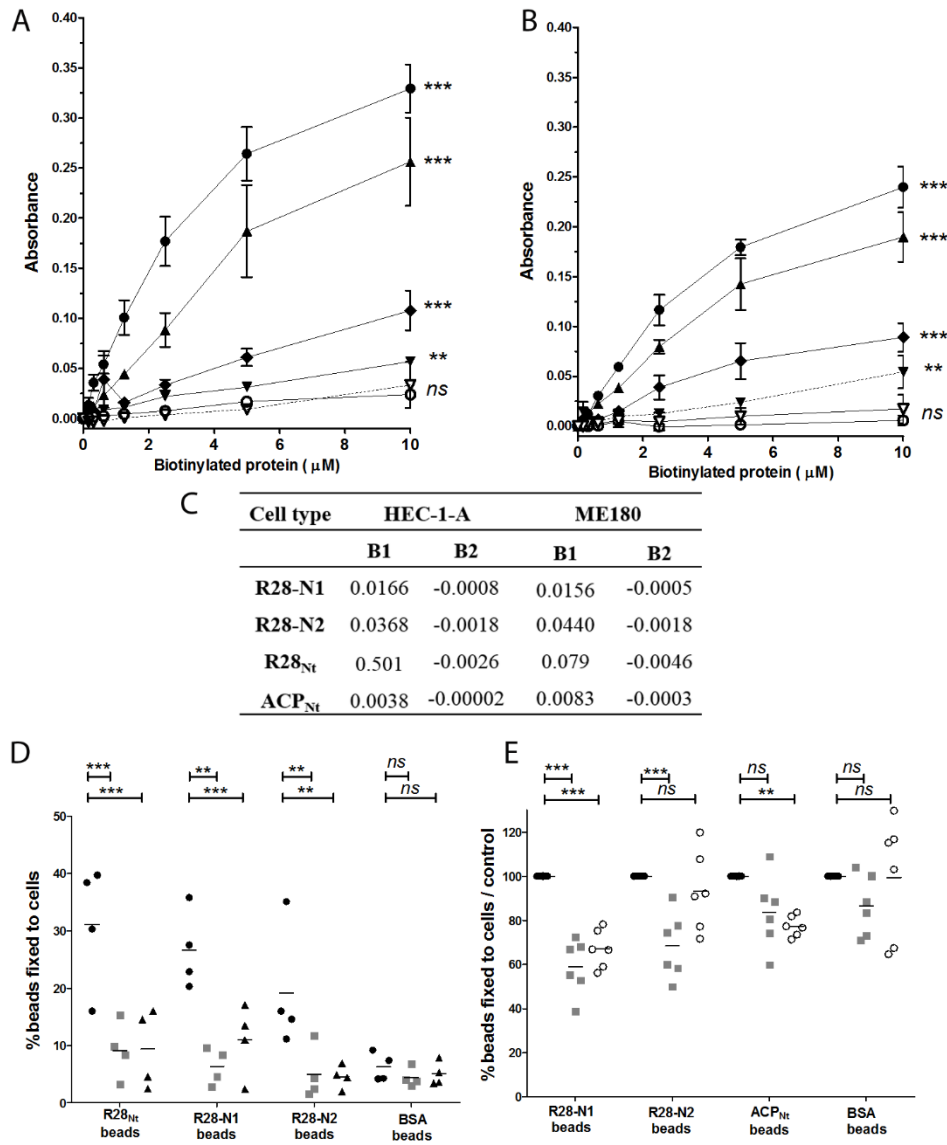
Figure 3. R28-N1, R28-N2 subdomains and ACP_Nt are sufficient to promote adhesion to HEC-1-A cells and compete differentially with each other binding. *A-B,* ELISA-based cell interaction assay with purified biotinylated R28_Nt (●), the subdomains R28-N1 (◆) and R28-N2 (▲); ACP_Nt (▼), Rib_Nt+2R (○) and BSA (▽). Peptide solutions were incubated with *A,* ME180 cells or *B,* HEC-1-A cell, washed, fixed with paraformaldehyde and bound biotinylated proteins were detected as in an ELISA. The experimental data for biotinylated R28_Nt and BSA are those used for Fig. 2*C* and 2*D*. The calculation for the absorbance is described in Experimental procedures. Error bars correspond to SEM of four independent experiments performed in duplicates. *C*, parameters of the non-linear fitting of figures *A* and *B*. *D-E,* Fluorescent (*D*) R28_Nt-, (*D, E*) R28-N1-, R28-N2-, (*E*) ACP_Nt- or (*D, E*) BSA-coated beads were incubated with HEC-1-A cells pre-incubated with 20 μM of recombinant R28-N1 (■); R28-N2, (▲), ACP_Nt (○) or not (●). **, p<0.005; ***, p<0.001; *ns*, not significant.

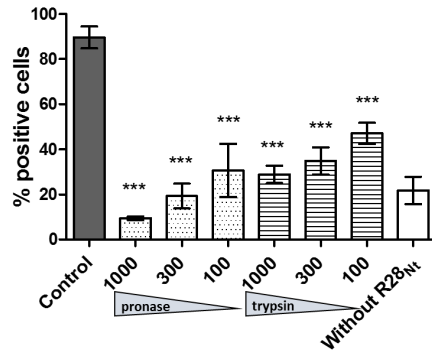**Figure 4. The R28$_{Nt}$ receptor is a cell surface protein.** Analysis of R28$_{Nt}$ receptor chemical nature. Treatments were applied to HEC-1-A cells in suspension prior to incubation with purified R28$_{Nt}$; bound R28$_{Nt}$ is immunostained and cells were analyzed by flow cytometry. Dark columns: untreated cells; clear column: unspecific labeling of cells without incubation of peptide. Statistical analysis was performed against the untreated condition. Error bars correspond to SEM of three independent experiments. Two-Way ANOVA: ***, p<0.001.

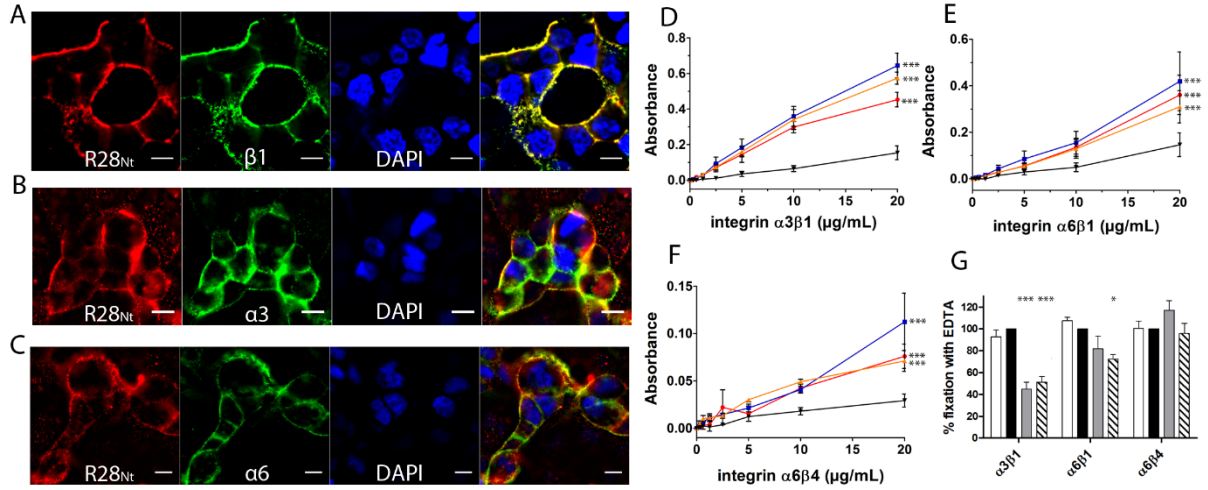**Figure 5. R28$_{Nt}$ and its subdomains interact with integrins α3β1, α6β1 and α6β4.** *A-C* Immunostaining of R28$_{Nt}$ and different integrin monomers on HEC-1-A cells: *A*, β1; *B*, α3; *C*, α6. R28$_{Nt}$, red; the specified integrin monomer, green; DAPI, blue; the right-hand side image, merge. Scale bar corresponds to 5 μm. *D-F* Assessment of integrin binding: *D*, α3β1; *E*, α6β1; *F*, α6β4 binding to R28$_{Nt}$, red; R28-N1, blue; R28-N2, orange; BSA, black, by ELISA. Error bars correspond to SEM of three to five independent experiments. Two-way-ANOVA at 20 μg/mL with Bonferroni post-tests against BSA. *G,* role of divalent cations in the binding of integrins to R28$_{Nt}$, PBS, clear column; PBS + EDTA, dark column; PBS + Mn^{2+}, grey column; PBS + Ca^{2+}, hatched column. One-Way ANOVA of four independent experiments. *, $p < 0.05$; ***, $p < 0.001$.

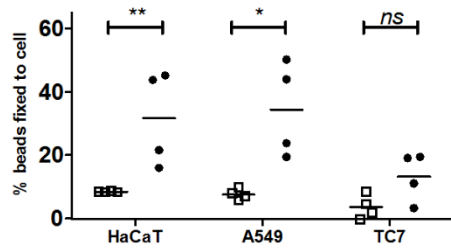**Figure 6. R28_{Nt} increases binding to different cell lines.** Binding of R28_{Nt}-, black circles; BSA-coated, empty squares, fluorescent beads to different epithelial cell types: HaCaT (keratinocytes), A549 (pulmonary) and TC7 (intestine). Four independent experiments performed in duplicates. Two-Way ANOVA. *, $p<0.05$ **; $p<0.005$; ns, not significant.