Pre-protein signature for full susceptibility to the co-translational translocation inhibitor cyclotriazadisulfonamide (CADA)

Victor Van Puyenbroeck[1], Eva Pauwels[1], Becky Provinciael[1], Thomas W. Bell[2], Dominique Schols[1], Kai-Uwe Kalies[3], Enno Hartmann[3] and Kurt Vermeire[1*]

[1]KU Leuven Department of Microbiology, Immunology and Transplantation, Rega Institute for Medical Research, Laboratory of Virology and Chemotherapy, B-3000 Leuven, Belgium

[2]Department of Chemistry, University of Nevada, Reno, NV 89557-0216, USA

[3]Centre for Structural and Cell Biology in Medicine, Institute of Biology, University of Lübeck, 23562 Lübeck, Germany

*__Correspondence__: Kurt Vermeire, KU Leuven, Rega Institute for Medical Research, Herestraat 49 – box 1043, 3000 Leuven, Belgium.

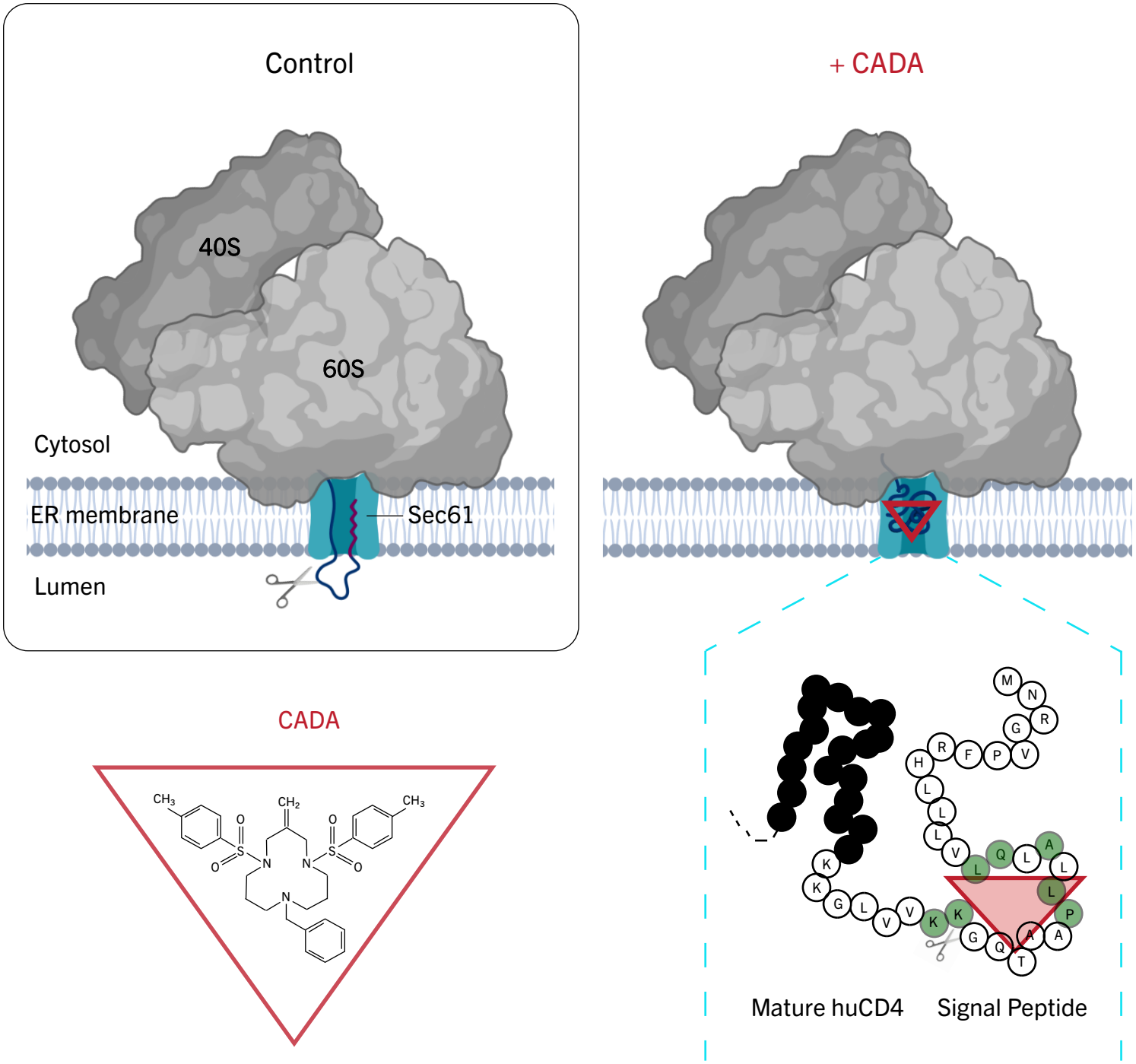
E-mail: kurt.vermeire@kuleuven.be

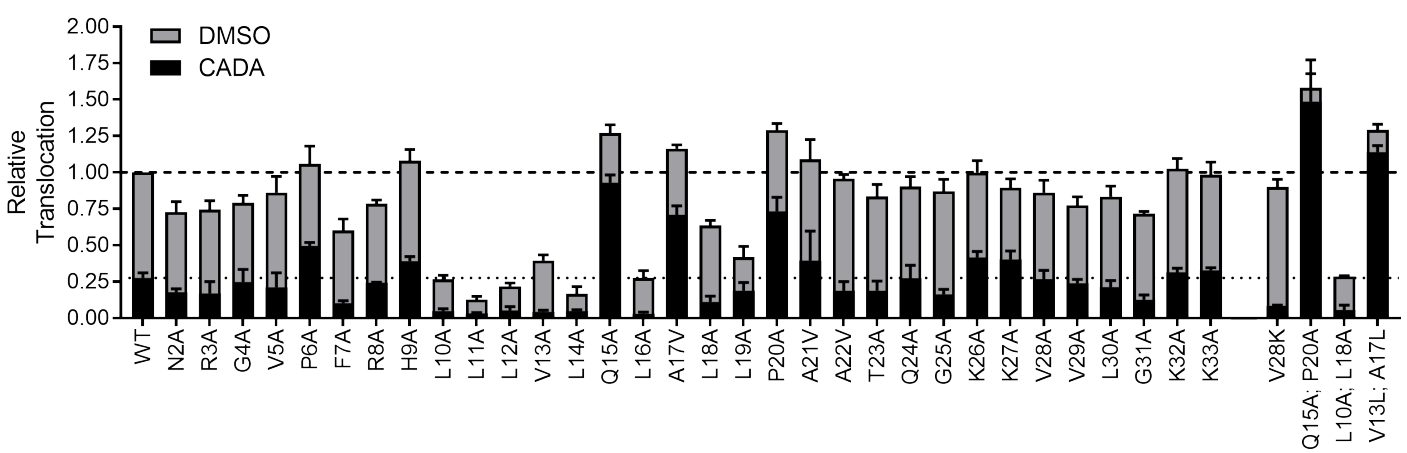
Running Title: Pre-protein signature for CADA sensitivity

Synopsis: Alanine scanning on the huCD4 signal peptide revealed that the general hydrophobicity of the hydrophobic (h)-domain of the signal peptide (SP) and positive charges in the mature protein are key elements that affect both the translocation efficiency of huCD4 and the sensitivity towards the small molecule translocation inhibitor CADA. In addition, unique amino acid pairs (L14/Q15 and L19/P20) in the SP hydrophobic core add specificity to the sensitivity signature for CADA.

Keywords: cyclotriazadisulfonamide, CADA, CD4, co-translational translocation, ER, small molecule inhibitor, signal peptide, alanine scanning

Acknowledgments: We like to thank Anita Camps, Joren Stroobants and Eric Fonteyn for excellent technical assistance and Anne Giraut for all her help to start up the alanine scan. We also thank Mark Marsh for critical reading of the manuscript. This work was partly supported by the KU Leuven grant no. PF/10/018.


Control
40S
60S
Cytosol
ER membrane
Sec61
Lumen
+ CADA
CADA
Mature huCD4
Signal Peptide


DMSO
CADA
Relative Translocation
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
WT
N2A
R3A
G4A
V5A
P6A
F7A
R8A
H9A
L10A
L11A
L12A
V13A
L14A
Q15A
L16A
A17V
L18A
L19A
P20A
A21V
A22V
T23A
Q24A
G25A
K26A
K27A
V28A
V29A
L30A
G31A
K32A
K33A
V28K
Q15A; P20A
L10A; L18A
V13L; A17L

ABSTRACT

Cyclotriazadisulfonamide (CADA) inhibits the co-translational translocation of human CD4 (huCD4) into the ER lumen in a signal peptide (SP)-dependent way. We propose that CADA binds the nascent huCD4 SP in a folded conformation within the translocon resembling a normally transitory state during translocation. Here, we used alanine scanning on the huCD4 SP to identify the signature for full susceptibility to CADA. In accordance with our previous work, we demonstrate that residues in the vicinity of the hydrophobic h-region are critical for sensitivity to CADA. In particular, exchanging Gln-15, Val-17 or Pro-20 in the huCD4 SP for Ala resulted in a resistant phenotype. Together with positively charged residues at the N-terminal portion of the mature protein, these residues mediate full susceptibility to the co-translational translocation inhibitory activity of CADA towards huCD4. In addition, sensitivity to CADA is inversely related to hydrophobicity in the huCD4 SP. *In vitro* translocation experiments confirmed that the general hydrophobicity of the h-domain and positive charges in the mature protein are key elements that affect both the translocation efficiency of huCD4 and the sensitivity towards CADA. Besides these two general SP parameters that determine the functionality of the signal sequence, unique amino acid pairs (L14/Q15 and L19/P20) in the SP hydrophobic core add specificity to the sensitivity signature for a co-translational translocation inhibitor.

1 INTRODUCTION

More than 30% of protein encoding genes in human tissue encode for secretory or membrane proteins.[1] Driven by specific targeting signals in their sequences, the majority of these proteins enter the secretory pathway in a Sec61 dependent manner in order to reach their final destination.[2-5] The successful translocation of proteins into the secretory pathway is essential for the proper functioning of cells, and disordered protein translocation in humans has been linked to several metabolic diseases including cancer.[6] The secretory pathway is highly conserved across all eukaryotes.[2,5,7] A key component of this pathway is the heterotrimeric Sec61 translocon composed of Sec61 α, β and γ monomers. The Sec61 translocon forms a gated aqueous channel in the endoplasmic reticulum (ER) membrane. Its hydrophilic interior allows the transport, or translocation, of pre-proteins across the lipid bilayer, while a lateral gate in the channel facilitates the integration of hydrophobic transmembrane protein segments into the lipid bilayer for the accommodation of integral membrane proteins.[8,9]

The secretory pathway operates two major modes of translocation. In post-translational translocation, completely synthesized polypeptide chains are targeted to, and subsequently translocated through the ER membrane. This translocation mechanism accommodates transport of small proteins and has been well studied in fungi and bacteria. Co-translational translocation on the other hand, the most common form of translocation seen for higher eukaryotes, directly couples protein synthesis to protein translocation through the Sec61 translocon.[3]

The co-translational translocation of pre-proteins from the cytosol into the ER lumen is a complex multistep process during which specific targeting signals, called signal peptides (SP) or signal sequences, play an essential role.[3,4,10,11] SPs are short hydrophobic sequences located at the N-terminal end of secretory or type I integral membrane proteins. These SPs, which are highly variable in sequence and primary structure, share several common features. They have a typical length of 15-30 residues and comprise a central hydrophobic core (h-region) flanked by a positively charged 'N-region' and a polar 'C-region' that contains the processing site for the signal peptidase.[12,13] When the SP emerges from the ribosome, the signal recognition particle (SRP) that is present in the cytosol,[14,15] binds via its SRP54 unit in the S domain to the SP while its Alu-domain interacts with the ribosome.[16,17] Upon this binding of SRP to the SP and ribosome, protein translation is stalled and the ribosome-nascent chain complex is targeted to the ER membrane through binding of SRP to its receptor.[18] Next, the nascent pre-protein chain is transferred to the Sec61 translocon, thereby reinitiating translation of the nascent chain and allowing translocation of the pre-protein across the ER membrane.[19] By coupling translation

to translocation, the energy that is generated by GTP hydrolysis during translation is used to drive the pre-protein through the translocon pore. Hence the name co-translational translocation.[20]

Over the past few years extensive research has been dedicated to finding inhibitors of ER co-translational translocation that block the biogenesis of secretory and membrane proteins.[21,22] Examples of compounds that interact with the Sec61 translocon are mycolactone,[23] apratoxin A,[24,25] and decatransin.[26] These inhibitors are reported to block ER translocation by binding to the Sec61 translocon, however they lack substrate specificity. The cyclic heptadepsipeptide HUN-7293 and its derivatives CAM741 and cotransin, have been described as compounds that potently inhibit co-translational translocation of a subset of secretory and type I transmembrane proteins in a signal peptide-discriminatory manner by binding to Sec61α.[27-29] In addition, the cotransin variant CT08 has been reported to be highly substrate-selective, but also to prevent the membrane integration of TNFα, a single-pass type II membrane protein with a non-cleavable signal anchor sequence, indicating that a cleavable signal sequence is not strictly required for cotransin sensitivity.[30]

We have previously shown that the small molecule cyclotriazadisulfonamide (CADA) selectively inhibits ER co-translational translocation of human CD4 (huCD4), through interaction with the huCD4 SP. CADA does not prevent the targeting of nascent huCD4 peptides to the translocon, but instead interferes with huCD4 biosynthesis at a post-targeting step.[31] From a proteomic study, sortilin was recently identified as a second substrate for CADA, although with lower susceptibility to the compound as compared to huCD4.[32]

Several research groups have proposed the “head-in-first” model, with the N-terminal SP being initially inserted towards the ER lumen.[33-35] In our previous report, we demonstrated that stalled nascent chains of huCD4 initially insert in an N_{lum}/C_{cyt} orientation before the final topology inversion of the SP takes place, in accordance with the “head-in-first” model. This inversion generates the looped conformation of the nascent protein chain inside the translocon, rendering the SP cleavage site accessible to the signal peptidase complex in the ER lumen. We hypothesize that CADA binds the huCD4 SP and prevents translocation of the pre-protein through inhibition of the SP topology inversion.[31] However, the exact mode of action of CADA remains to be unraveled. In our current study, we analyzed the contribution of each single residue in the huCD4 SP to the sensitivity to CADA by systematic mutagenesis, and aimed to identify important interaction sites for CADA within the huCD4 SP. Specific residues, mainly in the hydrophobic core of the huCD4 SP, have been identified as being critical for full sensitivity to CADA.

2 RESULTS

2.1 Mapping residues in the human CD4 signal peptide that contribute to CADA sensitivity

Previous work demonstrated that the amino-terminal region of huCD4 (denoted as $huCD4_{1-32}$) is essential for CADA sensitivity.[31] This region includes the signal peptide (25 residues) and seven residues of the mature huCD4 protein. In contrast to human CD4, cell surface expression of mouse CD4 (mCD4) is not affected by CADA. By exchanging the individual N-, h- and C-regions of the human signal peptide with the CADA-resistant mouse counterpart, the hydrophobic (h) region of the huCD4 signal peptide was identified as the primary requirement for CADA sensitivity, with some contribution from the C-region.[31]

In this study, we evaluated the involvement of each individual residue in $huCD4_{1-32}$ in the action of CADA (**Figure 1**), starting from an alanine scan mutagenesis. We performed this analysis on two different constructs: (i) full length (FL) huCD4 that is expressed on the cell surface, and (ii) yellow fluorescent protein (YFP) that carried the 32 amino-terminal residues of human CD4 ($huCD4_{1-32}$) at its N-terminus, and that - due to the presence of a signal peptide but lack of a membrane anchor - is secreted into the culture medium (**Figure 1A**). Alanine residues in the original sequence (i.e., Ala-17, Ala-21 and Ala-22) were replaced by valine residues (**Table 1**). Briefly, HEK293T cells were transiently transfected with the different plasmids and exposed to increasing doses of CADA to determine an IC_{50} value for protein down-modulation (**Figure 1C**), which allowed a quantitative comparison of CADA sensitivity for each mutant (**Table 1**).

CADA inhibited WT huCD4 expression with an IC_{50} of 0.54 µM, but was notably less effective on the inhibition of YFP secretion (IC_{50} of 1.27 µM) (**Figure 1C**). Alanine substitution of most amino acids within the N-terminal region did not result in major alteration in sensitivity to the compound (**Table 1**). These data are in line with our previous study with the human/mouse exchange mutants, which suggested that the N-region of the SP is not crucial for CADA's effect.[31] However, replacing the phenylalanine residue by an alanine at position 7 (F7A mutation) enhanced the sensitivity to CADA (**Table 1**). On the contrary, exchanging Arg-8 into alanine resulted in a clear decrease in sensitivity (8-fold and >39-fold for CD4 and YFP, respectively; **Figure 2A,B**). As the mouse ortholog also displays an arginine at the same position, our previous analysis could not point out the contribution of Arg-8 to CADA sensitivity.[31] The effect of this R8A replacement is due to the charge difference (loss of charge), because exchanging the arginine by the positively charged residue lysine, restored full sensitivity to CADA (**Figure 2A and Table S1**: R8K mutant). Introducing a negative charge at this position (R8E mutation) resulted in a signal peptide with reduced sensitivity to CADA (**Figure 2A and Table S1**: R8E mutant).

Analysis of the hydrophobic h-region of the SP identified three key amino acid residues that contributed to CADA sensitivity: Gln-15, Ala-17 and Pro-20 (**Table 1**). For both the full length huCD4 reporter as the secreted YFP construct, a Q15A substitution resulted in a complete loss of sensitivity, as evident from the dose-response curves (**Figure 2A,B**). Substituting Gln-15 with asparagine, a structurally related amino acid, partially restored CADA sensitivity (**Figure 2A and Table S1**, Q15N mutant), suggesting that an amide side chain is likely required at this position. Exchanging the alanine residue at position 17 for a valine resulted in a detectable decrease in sensitivity to CADA (**Figure 2 and Table 1**, A17V mutant). In addition, mutating Pro-20 into alanine clearly reduced the sensitivity to CADA, though not fully (**Figure 2**). As proline and glycine residues are known helix terminators, we exchanged Pro-20 for a Gly. The P20G mutation in the huCD4 SP fully restored CADA sensitivity (**Figure 2A and Table S1**, P20G mutant), indicating that an interruption of the hydrophobic alpha helix of the SP at this particular residue is key to maintain sensitivity to the drug. Furthermore, exchanging Pro-20 for a leucine residue revealed the additive effect of removing the alpha helix terminator on one hand and increasing the hydrophobicity on the other, a combination that rendered the SP completely resistant to CADA (**Figure 2**, P20L mutant).

As listed in **Table 1**, for most other residues in the h-region of the huCD4 SP, alanine substitution did not result in decreased sensitivity to CADA. On the contrary, one general observation was that exchanging leucine for an alanine residue mostly resulted in a slightly enhanced sensitivity to the compound (between 2- and 8-fold) (**Figure 3A,B**). Both for FL huCD4 and for huCD4$_{1\text{-}32}$-YFP, the alanine mutation at position 12 (Leu-12) enhanced CADA sensitivity most (8- and 5-fold, respectively; **Table 1**). Combining two of these substitutions (e.g. L10A and L18A) further enhanced the effect (**Table S1**, L10A;L18A mutant). Substituting leucine with an alanine residue reduces the hydrophobicity of the peptide side chain. Thus, lowering the hydrophobicity of the central core of the signal peptide generally increased the sensitivity to CADA in our cell-based flow cytometry assays. However, Leu-19 was the exception as evidenced by the reduced sensitivity of the L19A mutant in both reporter systems (**Figure 3C,D**, L19A mutant). On the other hand, increasing the hydrophobicity of the SP by an alanine into valine substitution had either no major effect (A21V and A22V), or resulted in reduced sensitivity to CADA (A17V) (**Figure 2**). Furthermore, an increase of the alpha helix's hydrophobicity by a double leucine insertion resulted in nearly full resistance to CADA (**Figure 3C and Table S1**, V13L;A17L mutant). Also, combining the Q15A and P20A mutations gave rise to a SP that no longer responded to CADA (**Figure 3C,D; Table S1**, Q15A;P20A mutant).

2.2 Charged residues in the mature protein region flanking the signal peptide contribute to CADA sensitivity

Previous work indicated that the first residues of the mature huCD4 protein add to CADA sensitivity.[31] To dissect the contribution of these N-terminal residues, mutagenesis was also performed on the first 7 amino acids of mature huCD4, i.e., residues 26 - 32 (**Table 1**). As shown in **Figure 4A,B**, alanine substitution of residues Lys-26 or Lys-27, directly C-terminal of the SP cleavage site, has a clear impact on CADA sensitivity. Exchange of the positively charged lysine residue at position 26 or 27 for a small, uncharged alanine residue resulted in reduction in sensitivity to CADA (**Figure 4A,B**), which was most evident in the YFP reporter system (IC_{50} > 50 µM, **Figure 4D and Table 1**, K26A mutant). To determine whether this effect was due to the loss of charge, we exchanged the lysine residue for a positively charged arginine. As displayed in **Figure 4A,B**, the K26R and K27R mutants kept full sensitivity to CADA. However, introduction of negative charges at these positions by replacing the lysines with glutamate residues (K26E and K27E) copied the alanine substitution phenotypes (**Figure 4A,B**). These data indicate that positive charges in the mature protein are crucial and suggest that electrostatic interactions are likely required for CADA's mechanism of action. Interestingly, introducing a third positive charge at the N-terminus of mature huCD4 by mutating Val-28 into a lysine residue increased the sensitivity to CADA by more than 3-fold, as evident from the smaller IC_{50} value (**Figure 4C and Table S1**, V28K mutant). Furthermore, charged residues down-stream of the SP cleavage site contribute to CADA sensitivity. When Lys-32 is mutated to alanine (K32A), a detectable reduction (~ 3-fold) in sensitivity was observed in the FL huCD4 system (**Figure 4C**), whereas the impact on secreted YFP was much more pronounced (**Figure 4D**; 14-fold increased IC_{50}). Of note, full-length huCD4 has an adjacent lysine at position 33 though, which is absent in the huCD4$_{1-32}$-YFP construct (**Figure 1A**), and this might compensate for the loss of one charge in the K32A mutant. In fact, the dose-response curve of the FL huCD4 K32A mutant (**Figure 4C**; red line) resembled that of the wild-type huCD4$_{1-32}$-YFP construct (**Figure 4D**; black line), with identical IC_{50} values (**Table 1**). Here, both constructs carry only a single positively charged lysine residue at position 32-33. Accordingly, replacing the adjacent lysine 33 in FL huCD4 with alanine (**Figure 4C and Table S1**; K33A) significantly reduced the CADA sensitivity of huCD4 (IC_{50} > 50 µM). Vice versa, inserting an extra lysine residue at the N-terminus of YFP resulted in an YFP reporter protein that phenotypically copied FL huCD4 in terms of CADA sensitivity (**Figure 4D**, insK33 mutant of secreted YFP; IC_{50} value: 0.44 µM). Thus, not only the amount of positively charged residues in this N-terminal region of the protein, but also the exact location of these lysine residues impact sensitivity to CADA. Our data also indicate that - in addition to the SP - a stretch of 8 amino acids (containing 4 lysine residues) at the N-terminus of the CD4 mature protein is required to preserve full susceptibility to CADA.

2.3 CADA differentially inhibits the co-translational translocation of huCD4 precursor protein with distinct signal peptide mutations

Because CADA directly affects protein translocation into the ER lumen during CD4 biosynthesis,[31] we next evaluated the co-translational translocation of huCD4 signal peptide mutants in a cell-free *in vitro* translation system. Transcripts of truncated huCD4 (i.e., the N-terminal D1D2 domains of huCD4 without a transmembrane anchor, and also deprived of sequons for N-glycosylation) were translated in the rabbit reticulocyte system in the absence or presence of ovine pancreatic microsomal membranes and exposed to different concentrations of CADA, as described elsewhere.[36] As shown in **Figure 5A,B**, translocation of huCD4 into the lumen of the microsomes (RM) was dose-dependently prevented by CADA for the WT construct, as determined by the qualitative ratio of processed SP-cleaved species (open arrowhead) to unprocessed intact pre-protein products (filled arrowhead). For the Q15A mutant, the impact of CADA on the co-translational translocation of huCD4 was significantly reduced. The P20A and the K26A mutants still responded to CADA, although to a lesser extent as compared to WT huCD4 (**Figure 5B**). In accordance with the flow cytometry analysis (**Figure 3C**), the Q15A;P20A double mutant exerted full resistance to CADA (**Figure 5A,B**).

2.4 Alanine mutation in the huCD4 signal peptide differentially affects translocation efficiency in a cell free translation system

Previous analysis suggested a possible variation in protein translocation levels between CADA sensitive and resistant substrates. We therefore performed a side-by-side comparison of the *in vitro* translocation efficiency for the different SP mutants. As summarized in **Figure 6A**, in the absence of CADA, a great number of alanine mutants of the huCD4 SP generally exerted lower translocation levels compared to the WT control. Furthermore, all mutants with a leucine into alanine substitution showed greatly reduced CD4 protein import into the ER lumen (**Figure 6A,B**), which was significant lower as compared to the WT control protein ($P < 0.005$, two-tailed unpaired t test with Welch's correction), indicating that mutants with reduced hydrophobicity of the SP become less functional in translocating the huCD4 protein. Consequently, these SP mutants exert higher sensitivity towards CADA (**Figure 6C**; black bars), which was most prominent and significant for the mutants L12A ($P = 0.0009$), L16A ($P = 0.0041$) and L18A ($P = 0.0017$). Remarkably, for the L14A and the L19A mutant, although a lower translocation efficiency was observed for the untreated controls as compared to WT (**Figure 6A**), CADA treatment was significantly less effective ($P = 0.006$ for L14A with 59% inhibition, and $P = 0.031$ for

L19A with 61% inhibition as compared to 75% for the WT protein). On the other hand, a significantly enhanced translocation as compared to the WT control ($P < 0.005$) was observed for those alanine mutants that were indicated from our initial alanine scan as being CADA resistant, such as the Q15A, A17V and P20A mutants (Figure 6A). The inhibitory effect of CADA on the protein translocation of these mutants was significantly reduced (Figure 6C; white bars). Furthermore, the Q15A; P20A double mutant exerted the highest translocation efficiency (Figure 6A) and responded significantly less to CADA ($P < 0.0001$), confirming the earlier observed resistance for this SP mutant (Figure 5). Regarding the lysine mutants, the general translocation efficiency of huCD4 was not affected (Figure 6A), but these K into A mutants showed significantly reduced sensitivity to CADA ($P = 0.0058$ for K26A and $P = 0.0048$ for K27A). Thus, alanine substitution of residues in the SP of huCD4 affected protein translocation into the ER in various degrees. Mutants with higher translocation efficiency as compared to WT huCD4 tend to respond less to CADA inhibition. By contrast, no strict correlation between reduced translocation efficiency and enhanced CADA sensitivity could be drawn.

2.5 Alanine mutation in the huCD4 signal peptide differentially affects cellular CD4 expression level

The above described alanine scan is based on transient transfection of cells with different constructs, encoding for either WT huCD4 or variants carrying mutations in the SP, followed by analysis of protein expression by flow cytometry. Next, we wanted to verify that the dissimilarity in CADA-sensitivity between the different mutants is a consequence of intrinsic compound resistance rather than variation in plasmid transfection efficiency and subsequent protein translation. Thus, additional experiments were performed with a construct encoding green fluorescent protein (GFP)-fused CD4 and red fluorescent protein (RFP) that are separated by the viral 2A sequence (Figure 7A), as described elsewhere.[37,38] At this P2A sequence, ribosomes skip formation of a peptide bond without interrupting translation elongation.[39] This way, each polycistronic mRNA encodes for two separated fluorescent proteins in equal amounts, with cytosolic RFP serving as an internal expression control that correlates with the number of transcripts, whereas the GFP signal being the measurement for huCD4 expression. In general, a similar down-modulating effect of CADA on huCD4-fused GFP was observed as with native huCD4 without the fluorescent tag (Figure 7B and 1C, respectively), indicating that fusion of GFP to the C-terminus of huCD4 does not change the susceptibility to CADA. Also, a comparable dose-response curve for CADA was obtained when cell surface CD4 was measured by means of a fluorescently labeled anti-CD4 antibody as when the GFP expression was quantified (Figure 7B). The impact of CADA on GFP levels was slightly stronger as on the level of CD4 measured by antibody staining, although the same huCD4tGFP fusion protein should be detected. However, GFP levels

represent the total population of CD4 molecules distributed over the different organelles and plasma membrane of the secretory pathway, whereas antibody-stained CD4 detects CD4 molecules at the cell surface. Also, in contrast to monomeric GFP, antibodies can crosslink several molecules making it more challenging to quantify the exact number of CD4 proteins with fluorescently tagged anti-CD4 antibodies.

For a selection of SP mutants, site-directed mutagenesis was performed to exchange the AA into an alanine (valine/leucine) in the SP of the huCD4-GFP-2A-RFP construct. The different SP mutants of GFP-tagged huCD4 responded to CADA in the same relative order as their non-tagged counterparts, both when GFP expression was analyzed (**Figure 7C and Table S2**) and when surface huCD4 was stained with a fluorescently labeled antibody (**Figure S1**). As summarized in **Figure 7D**, in the absence of CADA, a detectable increase (4 – 17 %) was measured in the cellular huCD4 expression of most SP mutants as compared to WT, suggesting a (slightly) more efficient protein translation and translocation of these huCD4 mutants. This effect was significant ($P < 0.05$, compared to WT control; two-tailed unpaired t test with Welch's correction) for the single mutants P6A, H9A, P20A and K26A, as well for both double mutants Q15A; P20A and V13L; A17L. However, the level of increase in huCD4 expression between the different SP mutants seems not to correlate with the degree of CADA resistance. For instance, both the CADA-resistant Q15A; P20A double mutant and the intermediate sensitive P20A mutant displayed a comparable significant increment in protein expression (14.3 and 12.5 %, respectively), irrespective of the level of enhanced translocation efficiency as quantified in **Figure 6A**. One possible explanation is that 24 hours after transfection the cells had already reached a steady level of huCD4 at the cell surface, so that higher translocation efficiency no longer contributed to actually raise the amount of huCD4 in a cellular context where different feed-back and check-point systems are at play. In contrast, the most striking effect was observed for the untreated L12A and the V28K mutants for which a 71 % and 61 %, respectively, decrease in huCD4 expression was measured as compared to FL huCD4 (**Figure 7D**), which was highly significant ($P < 0.0001$). In addition, for the CADA-sensitive L18A mutant, the reduction in huCD4 was only 10 %, whereas for the L19A mutant with reduced CADA susceptibility, the expression of huCD4 was decreased by 17 %. Notably, these mutants that expressed reduced GFP levels as compared to WT huCD4 were also identified as proteins with reduced translocation efficiency (**Figure 6A**), confirming that these huCD4 SP mutants do express lower absolute levels of huCD4. For L12A, L18A and L19A, the same relative order of protein expression was determined in both assays (**Figure 6A and 7D**). Although these results suggest a reduction in translation/translocation efficiency for huCD4 SP mutants that are more susceptible to CADA (**Table 1 and S1**), again no clear correlation could be drawn between the level of protein expression and the degree of CADA sensitivity (based on the IC_{50} values). However, in general, CADA resistance is observed for SP mutants with enhanced

cellular expression of huCD4, whereas SP mutants that are more sensitive to CADA appear to have lower cellular huCD4 protein expression levels.

3 DISCUSSION

Previously, the small-molecule macrocycle cyclotriazadisulfonamide (CADA) was identified as a selective huCD4 down-modulator. The compound inhibits huCD4 biogenesis by blocking the co-translational translocation of huCD4 into the ER, through direct binding to the cleavable N-terminal signal peptide of the huCD4 pre-protein.[31] Preliminary analysis of the huCD4 SP identified its hydrophobic h-region as the primary requirement for CADA sensitivity. In our current study, we determined the contribution of individual amino acid residues in the huCD4 SP towards CADA sensitivity by alanine scan mutagenesis.

In two independent reporter systems, we identified Gln-15, Ala-17 and Pro-20 in the hydrophobic core of the huCD4 SP as key amino acid residues that contributed to CADA sensitivity. The Q15A mutation especially resulted in resistance to CADA for the membrane anchored CD4 and soluble YFP reporter protein. Reduced effects of CADA on these mutant SPs was ascribed to decreased inhibition of the co-translational translocation of huCD4 in a cell free assay, indicating that a SP-related translocational event is involved. From additional data with a Q15N mutant, we concluded that the presence of a polar amide side chain in the hydrophobic alpha helix of the SP is likely required for interaction with CADA, presumably through one of the arenesulfonyl side arms of the compound, based on the electric dipole moment in the aromatic ring.[40] A glutamine in combination with a proline residue seemed to be the signature in huCD4 SP for full susceptibility to CADA, as substitution of both residues by an alanine resulted in full resistance to the drug in our cellular and cell free assays. Proline and glycine residues are known as helix terminators, while leucine has a high helical propensity.[41] Proline is present in the human SP but not in the resistant murine sequence, which contains a leucine residue at this position.[31] It might therefore play a critical role in the specificity of CADA for human over mouse CD4. Absence of this proline residue in the SP of mouse CD4 therefore implies that the alpha helix in this hydrophobic region offers a greater conformational stability that in turn might influence its interaction and positioning in the environment within the translocon. Of note, sortilin, the second substrate for CADA,[32] holds a SP that is quite distinct from the huCD4 SP in terms of length and amino acid composition, but it contains both a glutamine and two adjacent proline residues in the hydrophobic core of its SP. It would be interesting to further examine the role of these amino acids in the sensitivity to CADA in order to deduce a putative consensus sequence for full susceptibility to CADA.

We observed a clear impact on CADA sensitivity when the positively charged lysine residues K26 and K27, at the N-terminus of the mature huCD4 protein, were replaced with a neutral alanine residue. Charges in the mature protein domain directly downstream the signal peptide cleavage site contribute to the topology of the amino terminus during translocation. The effect of positively charged residues in bacterial membrane proteins was originally described by von Heijne,[42,43] where topology could be

predicted by a 'positive-inside rule': peptide regions with positively charged amino acid residues are located mainly on the cytoplasmic side ('inside') of the membrane. During co-translational translocation through the Sec61 translocon of uncleaved signal sequences (termed signal anchors), charges in the signal sequence were also shown to affect topology, similar to the 'positive inside rule'.[44,45] In addition, several studies propose the "head-in-first" model, with the N-terminal SP being inserted towards the ER lumen and subsequently inverted to form a loop in the translocon, ready for cleavage by the signal peptidase on the luminal side of the ER,[9,33,35] a model which we also proposed for huCD4.[31] The position of the lysine residues in huCD4 would favor an initial head-in-first insertion of the SP into the translocon channel, with the amino terminus facing the ER lumen and the C-terminus at the cytoplasmic side of the ER membrane (N_{lum}/C_{cyt}). Replacing the positively charged Lys-26 or Lys-27 residues with a neutral alanine could shift the charge balance towards the N-terminus where the Arg-3 and Arg-8 residues help to retain the N-terminus of the SP at the cytoplasmic side of the translocon. This could force the SP to insert directly in the looped conformation, bypassing the crucial topology inversion on which CADA is believed to operate. Furthermore, the observation that an additional lysine residue at the N-terminus of the mature huCD4 protein (V28K mutant) greatly improved sensitivity of huCD4 to CADA suggests that the net positive charge of the N-terminus of the mature protein determines if the protein becomes a substrate for CADA. Additionally, charges at the N-terminal domain of the SP seem to contribute to the sensitivity to CADA as well. Positively charged residues at the N region of the SP are very important for the inversion process in the head-in-first model. Removal of the most N-terminal positioned arginine residue (Arg-3) would favor enhanced head-in-first insertion of the SP according to the 'positive inside rule'. Expectedly, this R3A mutant exerted a slightly enhanced sensitivity to CADA, supporting our hypothesis that CADA inhibits the SP topology inversion. However, mutating the more down-stream arginine (Arg-8) resulted in reduced sensitivity towards CADA. As this residue is located at the edge of the hydrophobic core of the SP, introducing an alanine residue at this position (R8A mutant) could simply enhance the total hydrophobicity of the h-region that overcomes the loss in charge. Thus, not only the amount of positively charged residues in this N-terminal region of the protein, but also the exact location of the charges influences CADA sensitivity.

Our study demonstrates that sensitivity to CADA does not reside in the signal sequence alone, but that the N-terminus of the mature protein is important as well. It appears reasonable to assume that the primary sequence of signal peptides has evolved to work in cooperation with the mature protein sequence, as suggested by Kim *et al*.[46] Signal peptide cleavage for example, requires the formation of a looped topology inside the translocon. This brings the signal peptide and the downstream mature protein sequence into close proximity. As our current model of CADA's mode of action also proposes

a looped (and/or folded) formation of the SP and mature part of the substrate in the translocon, the combination of both protein parts will determine the final sensitivity to CADA as the compound might stabilize one of the specific conformations formed by the SP and mature protein region. Furthermore, the discovery of critical residues in the mature protein domain for full sensitivity to CADA has major consequences for the future design of an efficient SP reporter model. A signal peptide library could be constructed for high-throughput screening of CADA compounds or other signal peptide-dependent translocation inhibitors such as cotransin.[28] However, we have shown here that peptides in such a library should also include sufficient residues from the mature protein region to represent the natural substrates in a reliable form.

Sensitivity to CADA also largely depends on the hydrophobicity of the huCD4 signal peptide: mutations that reduce the hydrophobicity of the h-region generally increased the effect of CADA (lower IC_{50} values for CD4 down-modulation), while an increase in hydrophobicity provided resistance to CADA's inhibitory effect. Signal peptides with increased hydrophobicity are thus able to escape CADA-induced translocation inhibition. Our data from the cell free translation/translocation assay demonstrated the reduced translocation efficiency for the leucine into alanine mutants. The lower hydrophobicity of the alanine mutants could slow down the insertion and co-translational translocation of the nascent chain. This would provide CADA more time to stabilize the (hypothesized) folded state of the signal peptide to block translocation. In general, huCD4 mutants that were translocated more efficiently as compared to the WT control exerted lower sensitivity to CADA. In fact, the highest translocation level was noted for the Q15A; P20A mutant, which was also the mutant that exerted full resistance to CADA. Remarkably, bovine pre-prolactin and mouse CD4, both completely resistant to CADA,[31] are proteins that translocated more efficiently as compared to huCD4 (data not shown). However, comparison of two adjacent L18A and L19A mutants indicate that sensitivity to CADA is not just a matter of hydrophobicity and translocation efficiency. Whereas both mutants do translocate huCD4 with lower efficiency as compared to the WT control, the relative inhibition of CADA on these mutants was completely opposite (i.e., enhanced and reduced sensitivity to CADA for L18A and L19A, respectively). Interestingly, L14A and L19A, the only leucine into alanine mutants with reduced translocation efficiency and reduced CADA sensitivity, are located adjacent (one residue upstream) to the critical Q15 and P20, which might suggest specific connecting sites within the SP for interaction with CADA.

Comparable effects regarding SP hydrophobicity and translocation inhibition have been described for the related cotransin compounds. Reducing the hydrophobicity of VCAM-1 and VEGF-1 SP's h-region increases the inhibitory effect of CAM741, and vice-versa, enhanced hydrophobicity relates to resistance to CAM741,[47,48] suggesting that hydrophobicity of the cleavable signal peptide is a key parameter for cotransin sensitivity. However, the discovery of TNFα as a cotransin target disproved

this hypothesis, because TNFα contains a signal anchor sequence instead of a cleavable SP.[30] Signal anchors need to span the entire lipid bilayer and are thus significantly more hydrophobic than most cleavable signal peptides. Based on the huCD4 SP hydrophobicity mutants our current data suggest that CAM741/cotransin and CADA might share some features in how they prevent the translocation of their substrates, despite having a very different structure.

Signal peptides show no homology in their primary sequence, but they share a common structure[12] to perform two main functions: serving as a universal signal to target the ribosome nascent chain complexes to the ER membrane, and initiating the translocation across the ER membrane.[10,14] There is increasing evidence that sequence variation among signal sequences affects the efficiency of protein targeting to the translocon, the level of protein gating, the translocation across the ER membrane, and also the efficiency of signal peptide cleavage by the signal peptidase.[46,49] The primary sequence of each SP will thus dictate the final functionality of the SP and subsequently determine the efficiency and speed of the "entry" of the substrate into the secretory pathway. A recent study with the Xbp1 arrest peptide by Kriegler *et al.* demonstrated that efficient (or 'strong') signal peptides experience a two-phase pulling event whereby the nascent chain is pulled from the ribosome during its translocation.[35] The first (weaker) pulling force exerted on the SP is during early (head-in-first) engagement of the nascent chain with the translocon channel, whereas the second (stronger) pulling phase represents the SP inversion inside the translocon. Interestingly, the force profile was highly dependent on the hydrophobicity of the h-domain of the SP and on the positive charge of the N-region of the SP.[35] As we hypothesize that CADA stabilizes an intermediate folded state of the SP during the topology inversion, the compound/nascent chain/translocon complex should be able to resist that second pulling force to prevent further protein translocation. In that perspective, it would be very interesting to measure and analyze the pulling force profile of different huCD4 SP mutants and other potential substrates in order to see if it correlates with the level of susceptibility to CADA.

Altogether, our data point to hydrophobicity and charges in the SP of huCD4 as two key elements that influence full susceptibility to CADA. In addition to these general parameters of the SP that determine the functionality of the signal sequence, unique amino acid pairs (L14/Q15 and L19/P20) in the hydrophobic core of the SP add specificity to the sensitivity signature. Future experiments to visualize the CADA-treated huCD4 SP in the translocon by means of Cryo-EM would be of great help to improve our understanding of blocking co-translational translocation with small molecule inhibitors at the molecular level.

4 MATERIALS AND METHODS

4.1 Compound

Cyclotriazadisulfonamide (CADA) hydrochloride was synthesized as described elsewhere,[50] dissolved in dimethyl sulfoxide (DMSO) and stored at room temperature.

4.2 Plasmids and mutagenesis

The pcDNA3 expression vector (Invitrogen) encoding wild-type human CD4 (pcCD4), denoted as construct FL hCD4, has been described previously.[31] The coding sequence for residues 1-32 from the human CD4 pre-protein was fused to the sequence of yellow fluorescent protein (YFP) by PCR and subcloned into the pcDNA3.1D vector to generate the pc$CD4_{1-32}$-YFP plasmid, denoted as construct h$CD4_{1-32}$-YFP_{2-251}. The pEGFP-P2A-RFP expression vector (pEGFP-N1 Clontech backbone) was a kind gift from Dr. Ramanujan Hegde.[37] The sequence for enhanced GFP (eGFP) was replaced by that of turbo GFP (tGFP) and wild-type hCD4 was fused to the N-terminus of tGFP as depicted in Figure 7A. Site-directed mutagenesis of constructs was performed with the QuikChange II kit (Stratagene), Q5 site-directed mutagenesis kit (New England Biolabs) or NEBuilder HiFi DNA assembly kit (New England Biolabs), following the manufacturer's instructions. Plasmid DNA was isolated using the Wizard Plus SV system (Promega), supplemented with an endotoxin removal wash. The concentration of all constructs was determined with a NanoDrop 1000 spectrophotometer and sequences were confirmed by automated capillary Sanger sequencing (Macrogen Europe).

4.3 Transient transfection and flow cytometry

HEK293T cells were cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% (v/v) fetal bovine serum and 10 mM HEPES. For transfection, cells were seeded in DMEM supplemented with 10% serum, 1 mM sodium pyruvate (Invitrogen) and 0.075% (m/v) $NaHCO_3$ (Invitrogen) and allowed to adhere overnight. Lipofectamine 2000 and Lipofectamine LTX (Invitrogen) were used for the transfection of plasmid DNA, according to the manufacturer's instructions. CADA was added 6 hours post transfection and cells were grown for one or two days in the presence of the compound (as indicated in the legend of the figures).

CD4 cell surface expression was determined by immunostaining with Fluorescein isothiocyanate (FITC) or Allophycocyanin (APC)-labeled anti-human CD4 [clone SK3] (BioLegend). To measure YFP expression, the supernatant of pc$CD4_{1-32}$-YFP-transfected HEK293T cells was mixed with GFP-Trap-M magnetic beads (Chromotek) for 1 hour. Beads were then washed with phosphate-buffered saline before measurement of the YFP signal by flow cytometry. Levels of cellular tGFP and RFP were

simultaneously analyzed using flow cytometry. Fluorescence data were collected on either a BD FACSCalibur flow cytometer (Beckton Dickinson) equipped with CellQuest v3.3 software, a BD FACS Canto II or a BD FACS Fortessa flow cytometer (Beckton Dickinson) with BD FACSDiva 8.0.1 software. All data were analyzed in FlowJo X v10.

4.4 Cell-free *in vitro* translation and translocation

The Qiagen EasyXpress linear template kit was used to generate full length cDNAs. Wild-type full-length huCD4 contains glycosylation sites in the extracellular immunoglobulin-like domains D3 and D4.[51] To simplify analysis of the translated products, we generated truncated huCD4 (comprising the N-terminal D1 and D2 domains) using PCR. PCR products were purified and transcribed *in vitro* using T7 RNA polymerase (RiboMAX system, Promega). These mRNAs contain a STOP codon after the second immunoglobulin domain and encode C-terminally truncated huCD4D1D2 (i.e., the N-terminal D1D2 domains of huCD4 without the glycosylation sites and the transmembrane anchor). All transcripts were translated in rabbit reticulocyte lysate (Promega) in the presence of L-^{35}S-methionine (Perkin Elmer). Translations were performed at 30°C in the presence or absence of ovine pancreatic microsomes and CADA. Proteinase K protection assays were performed on ice for 30 min and quenched with phenylmethylsulfonyl fluoride. Samples were washed with low-salt buffer (80 mM KOAc, 2 mM $Mg(OAc)_2$, 50 mM HEPES pH 7.6) and radiolabeled proteins were isolated by centrifugation for 10 minutes at 21,382×*g* and 4°C (Hettich 200R centrifuge with 2424-B rotor). The proteins were then separated with SDS-PAGE and detected by phosphor imaging (Cyclone Plus storage phosphor system, Perkin Elmer).

4.5 Statistical analysis

Statistical analyses were performed using Graph Pad Prism version 7.04 for Windows (Graph Pad Software; www.graphpad.com). Two-tailed unpaired t test with Welch's correction was used for the analysis of the *in vitro* translocation data and the GFP:RFP ratio for huCD4 SP mutants of the huCD4tGFP-2A-RFP constructs. Observed differences were regarded as significant if the calculated *P*-values were ≤ 0.05. **P* < 0.05; ***P* < 0.01; ****P* < 0.001.

ACKNOWLEDGMENTS

We like to thank Anita Camps, Joren Stroobants and Eric Fonteyn for excellent technical assistance and Anne Giraut for all her help to start up the alanine scan. We also thank Mark Marsh for critical reading of the manuscript. This work was partly supported by the KU Leuven grant no. PF/10/018.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

AUTHOR CONTRIBUTIONS

V.V.P, E.H. and K.V. conceptualized and designed the project. V.V.P., E.P. and B.P. conducted the experiments. T.W.B. and D.S. provided the reagents. V.V.P., E.P. and K.V. analyzed the data. V.V.P., E.P. and K.V. prepared the manuscript. V.V.P., E.P., T.W.B., D.S., K.K., E.H., and K.V. wrote the paper and finalized the manuscript. All authors read and approved the final manuscript.

ORCID

Kurt Vermeire http://orcid.org/0000-0003-1123-1907.

References

1. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347(6220):1260419.

2. Wickner W, Schekman R. Protein translocation across biological membranes. *Science*. 2005;310(5753):1452-1456.

3. Rapoport TA. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature*. 2007;450(7170):663-669.

4. Hegde RS, Kang SW. The concept of translocational regulation. *J Cell Biol*. 2008;182(2):225-232.

5. Aviram N, Schuldiner M. Targeting and translocation of proteins to the endoplasmic reticulum at a glance. *J Cell Sci*. 2017;130(24):4079-4085.

6. Lang S, Pfeffer S, Lee PH, Cavalie A, Helms V, Forster F, et al. An Update on Sec61 Channel Functions, Mechanisms, and Related Diseases. *Front Physiol*. 2017;8:1-22.

7. Pfeffer S, Dudek J, Zimmermann R, Forster F. Organization of the native ribosome-translocon complex at the mammalian endoplasmic reticulum membrane. *Biochim Biophys Acta*. 2016;1860(10):2122-2129.

8. van den Berg B, Clemons WM, Jr., Collinson I, Modis Y, Hartmann E, Harrison SC, et al. X-ray structure of a protein-conducting channel. *Nature*. 2004;427(6969):36-44.

9. Voorhees RM, Hegde RS. Structure of the Sec61 channel opened by a signal sequence. *Science*. 2016;351(6268):88-91.

10. Blobel G, Dobberstein B. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J Cell Biol*. 1975;67(3):835-851.

11. Walter P, Gilmore R, Blobel G. Protein translocation across the endoplasmic reticulum. *Cell*. 1984;38(1):5-8.

12. von Heijne G. Signal sequences. The limits of variation. *J Mol Biol*. 1985;184(1):99-105.

13. Hegde RS, Bernstein HD. The surprising complexity of signal sequences. *Trends Biochem Sci*. 2006;31(10):563-571.

14. Walter P, Blobel G. Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. *J Cell Biol*. 1981;91(2 Pt 1):551-556.

15. Keenan RJ, Freymann DM, Stroud RM, Walter P. The signal recognition particle. *Annu Rev Biochem*. 2001;70:755-775.

16. Halic M, Beckmann R. The signal recognition particle and its interactions during protein targeting. *Curr Opin Struct Biol*. 2005;15(1):116-125.

17. Voorhees RM, Hegde RS. Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. *Elife*. 2015;4:e07975.

18. Gilmore R, Walter P, Blobel G. Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. *J Cell Biol*. 1982;95(2):470-477.

19. Park E, Rapoport TA. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu Rev Biophys*. 2012;41:21-40.

20. Saraogi I, Shan SO. Molecular mechanism of co-translational protein targeting by the signal recognition particle. *Traffic*. 2011;12(5):535-542.

21. Kalies KU, Romisch K. Inhibitors of Protein Translocation Across the ER Membrane. *Traffic*. 2015;16(10):1027-1038.

22. Van Puyenbroeck V, Vermeire K. Inhibitors of protein translocation across membranes of the secretory pathway: novel antimicrobial and anticancer agents. *Cell Mol Life Sci*. 2018;75(9):1541-1558.

23. Hall BS, Hill K, McKenna M, Ogbechi J, High S, Willis AE, et al. The pathogenic mechanism of the Mycobacterium ulcerans virulence factor, mycolactone, depends on blockade of protein translocation into the ER. *PLoS Pathog*. 2014;10(4):e1004061.

24. Paatero AO, Kellosalo J, Dunyak BM, Almaliti J, Gestwicki JE, Gerwick WH, et al. Apratoxin Kills Cells by Direct Blockade of the Sec61 Protein Translocation Channel. *Cell Chem Biol*. 2016;23(5):561-566.

25. Huang KC, Chen Z, Jiang Y, Akare S, Kolber-Simonds D, Condon K, et al. Apratoxin A Shows Novel Pancreas-Targeting Activity through the Binding of Sec 61. *Mol Cancer Ther*. 2016;15(6):1208-1216.

26. Junne T, Wong J, Studer C, Aust T, Bauer BW, Beibel M, et al. Decatransin, a new natural product inhibiting protein translocation at the Sec61/SecYEG translocon. *J Cell Sci*. 2015;128(6):1217-1229.

27. Besemer J, Harant H, Wang S, Oberhauser B, Marquardt K, Foster CA, et al. Selective inhibition of cotranslational translocation of vascular cell adhesion molecule 1. *Nature*. 2005;436(7048):290-293.

28. Garrison JL, Kunkel EJ, Hegde RS, Taunton J. A substrate-specific inhibitor of protein translocation into the endoplasmic reticulum. *Nature*. 2005;436(7048):285-289.

29. Mackinnon AL, Paavilainen VO, Sharma A, Hegde RS, Taunton J. An allosteric Sec61 inhibitor traps nascent transmembrane helices at the lateral gate. *Elife*. 2014;3:e01483.

30. Maifeld SV, MacKinnon AL, Garrison JL, Sharma A, Kunkel EJ, Hegde RS, et al. Secretory protein profiling reveals TNF-alpha inactivation by selective and promiscuous Sec61 modulators. *Chem Biol*. 2011;18(9):1082-1088.

31. Vermeire K, Bell TW, Van Puyenbroeck V, Giraut A, Noppen S, Liekens S, et al. Signal peptide-binding drug as a selective inhibitor of co-translational protein translocation. *PLoS Biol*. 2014;12(12):e1002011.

32. Van Puyenbroeck V, Claeys E, Schols D, Bell TW, Vermeire K. A Proteomic Survey Indicates Sortilin as a Secondary Substrate of the ER Translocation Inhibitor Cyclotriazadisulfonamide (CADA). *Mol Cell Proteomics*. 2017;16(2):157-167.

33. Goder V, Spiess M. Molecular mechanism of signal sequence orientation in the endoplasmic reticulum. *EMBO J*. 2003;22(14):3645-3653.

34. Voorhees RM, Hegde RS. Toward a structural understanding of co-translational protein translocation. *Curr Opin Cell Biol*. 2016;41:91-99.

35. Kriegler T, Magoulopoulou A, Amate Marchal R, Hessa T. Measuring Endoplasmic Reticulum Signal Sequences Translocation Efficiency Using the Xbp1 Arrest Peptide. *Cell Chem Biol*. 2018;25(7):880-890.

36. Vermeire K, Allan S, Provinciael B, Hartmann E, Kalies KU. Ribonuclease-neutralized pancreatic microsomal membranes from livestock for in vitro co-translational protein translocation. *Anal Biochem*. 2015;484:102-104.

37. Juszkiewicz S, Hegde RS. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol Cell*. 2017;65(4):743-750.

38. Lin YJ, Huang LH, Huang CT. Enhancement of heterologous gene expression in Flammulina velutipes using polycistronic vectors containing a viral 2A cleavage sequence. *PLoS One*. 2013;8(3):e59099.

39. de Felipe P, Hughes LE, Ryan MD, Brown JD. Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. *J Biol Chem*. 2003;278(13):11441-11448.

40. Chawla R, Van Puyenbroeck V, Pflug NC, Sama A, Ali R, Schols D, et al. Tuning Side Arm Electronics in Unsymmetrical Cyclotriazadisulfonamide (CADA) Endoplasmic Reticulum (ER) Translocation Inhibitors to Improve their Human Cluster of Differentiation 4 (CD4) Receptor Down-Modulating Potencies. *J Med Chem*. 2016;59(6):2633-2647.

41. Chou PY, Fasman GD. Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry-Us*. 1974;13(2):211-222.

42. von Heijne G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol*. 1992;225(2):487-494.

43. von Heijne G. Net N-C charge imbalance may be important for signal sequence function in bacteria. *J Mol Biol*. 1986;192(2):287-290.

44. Kocik L, Junne T, Spiess M. Orientation of internal signal-anchor sequences at the Sec61 translocon. *J Mol Biol*. 2012;424(5):368-378.

45. Beltzer JP, Fiedler K, Fuhrer C, Geffen I, Handschin C, Wessels HP, et al. Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. *J Biol Chem*. 1991;266(2):973-978.

46. Kim SJ, Mitra D, Salerno JR, Hegde RS. Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev Cell*. 2002;2(2):207-217.

47. Harant H, Lettner N, Hofer L, Oberhauser B, de Vries JE, Lindley IJ. The translocation inhibitor CAM741 interferes with vascular cell adhesion molecule 1 signal peptide insertion at the translocon. *J Biol Chem*. 2006;281(41):30492-30502.

48. Harant H, Wolff B, Schreiner EP, Oberhauser B, Hofer L, Lettner N, et al. Inhibition of vascular endothelial growth factor cotranslational translocation by the cyclopeptolide CAM741. *Mol Pharmacol*. 2007;71(6):1657-1665.

49. Levine CG, Mitra D, Sharma A, Smith CL, Hegde RS. The efficiency of protein compartmentalization into the secretory pathway. *Mol Biol Cell*. 2005;16(1):279-291.

50. Bell TW, Anugu S, Bailey P, Catalano VJ, Dey K, Drew MG, et al. Synthesis and structure-activity relationship studies of CD4 down-modulating cyclotriazadisulfonamide (CADA) analogues. *J Med Chem*. 2006;49(4):1291-1312.

51. Spellman MW, Leonard CK, Basa LJ, Gelineo I, van Halbeek H. Carbohydrate structures of recombinant soluble human CD4 expressed in Chinese hamster ovary cells. *Biochemistry-Us*. 1991;30(9):2395-2406.

Table 1. Differential CADA sensitivity of alanine scan mutants.

	Signal peptide	Cellular huCD4[†]					Secreted YFP[#]			
	N h C	IC_{50} (µM)[‡]					IC_{50} (µM)			
	1 5 10 15 20 25 32	Low	**Mean**	High[§]	*FC*[¶]		Low	**Mean**	High	*FC*
WT	MNRGVPFRH**LLLVLQLALLP**AATQG-KKVVLGK	0.47	**0.54**	0.62	*1.00*		1.15	**1.27**	1.42	*1.00*
N2A	M**A**RGVPFRHLLLVLQLALLPAATQG-KKVVLGK	0.55	**0.73**	1.12	*+1.4*		0.24	**0.39**	0.66	*-3.3*
R3A	MN**A**GVPFRHLLLVLQLALLPAATQG-KKVVLGK	0.34	**0.47**	0.67	*-1.1*		1.82	**1.99**	2.20	*+1.6*
G4A	MNR**A**VPFRHLLLVLQLALLPAATQG-KKVVLGK	0.26	**0.35**	0.47	*-1.5*		0.66	**0.91**	1.35	*-1.4*
V5A	MNRG**A**PFRHLLLVLQLALLPAATQG-KKVVLGK	0.20	**0.29**	0.40	*-1.9*		0.23	**0.33**	0.49	*-3.8*
P6A	MNRGV**A**FRHLLLVLQLALLPAATQG-KKVVLGK	0.53	**0.96**	2.06	*+1.8*		0.20	**0.55**	1.75	*-2.3*
F7A	MNRGVP**A**RHLLLVLQLALLPAATQG-KKVVLGK	0.11	**0.13**	0.17	*-4.2*		0.31	**0.46**	0.71	*-2.8*
R8A	MNRGVPF**A**HLLLVLQLALLPAATQG-KKVVLGK	2.63	**4.20**	8.68	*+7.8*		>50	**>50**	>50	*>+39*
H9A	MNRGVPFR**A**LLLVLQLALLPAATQG-KKVVLGK	0.20	**0.62**	2.28	*+1.1*		1.57	**2.13**	3.00	*+1.7*
L10A	MNRGVPFRH**A**LLVLQLALLPAATQG-KKVVLGK	0.16	**0.20**	0.25	*-2.7*		0.45	**0.60**	0.83	*-2.1*
L11A	MNRGVPFRHL**A**LVLQLALLPAATQG-KKVVLGK	0.10	**0.12**	0.14	*-4.5*		0.33	**0.38**	0.45	*-3.3*
L12A	MNRGVPFRHLL**A**VLQLALLPAATQG-KKVVLGK	0.03	**0.07**	0.11	*-7.7*		0.20	**0.24**	0.28	*-5.3*
V13A	MNRGVPFRHLLL**A**LQLALLPAATQG-KKVVLGK	0.12	**0.14**	0.18	*-3.9*		0.41	**0.46**	0.51	*-2.8*
L14A	MNRGVPFRHLLLV**A**QLALLPAATQG-KKVVLGK	0.37	**0.46**	0.59	*-1.2*		0.83	**1.12**	1.53	*-1.1*
Q15A	MNRGVPFRHLLLVL**A**LALLPAATQG-KKVVLGK	>50	**>50**	>50	*>+93*		>50	**>50**	>50	*>+39*
L16A	MNRGVPFRHLLLVLQ**A**ALLPAATQG-KKVVLGK	0.10	**0.11**	0.13	*-4.9*		0.34	**0.40**	0.47	*-3.2*
A17V	MNRGVPFRHLLLVLQL**V**LLPAATQG-KKVVLGK	1.00	**1.33**	1.76	*+2.5*		>50	**>50**	>50	*>+39*
L18A	MNRGVPFRHLLLVLQLA**A**LPAATQG-KKVVLGK	0.09	**0.10**	0.12	*-5.4*		0.59	**0.67**	0.78	*-1.9*
L19A	MNRGVPFRHLLLVLQLAL**A**PAATQG-KKVVLGK	0.89	**1.04**	1.23	*+1.9*		3.12	**4.57**	7.37	*+3.6*
P20A	MNRGVPFRHLLLVLQLALL**A**AATQG-KKVVLGK	0.86	**2.35**	>50	*+4.4*		>50	**>50**	>50	*>+39*
A21V	MNRGVPFRHLLLVLQLALLP**V**ATQG-KKVVLGK	0.48	**0.87**	1.73	*+1.6*		0.63	**0.77**	1.00	*-1.6*
A22V	MNRGVPFRHLLLVLQLALLPA**V**TQG-KKVVLGK	0.46	**0.79**	1.62	*+1.5*		1.22	**1.58**	2.10	*+1.2*
T23A	MNRGVPFRHLLLVLQLALLPAA**A**QG-KKVVLGK	0.63	**0.87**	1.29	*+1.6*		2.20	**2.36**	2.56	*+1.9*
Q24A	MNRGVPFRHLLLVLQLALLPAAT**A**G-KKVVLGK	0.89	**1.22**	1.54	*+2.3*		3.64	**4.22**	4.98	*+3.3*
G25A	MNRGVPFRHLLLVLQLALLPAATQ**A**-KKVVLGK	0.52	**0.61**	0.74	*+1.1*		0.18	**0.35**	0.66	*-3.6*
	Mature protein									
K26A	MNRGVPFRHLLLVLQLALLPAATQG-**A**KVVLGK	0.89	**1.44**	2.02	*+2.7*		>50	**>50**	>50	*>+39*
K27A	MNRGVPFRHLLLVLQLALLPAATQG-K**A**VVLGK	0.84	**1.30**	1.98	*+2.4*		>50	**>50**	>50	*>+39*
V28A	MNRGVPFRHLLLVLQLALLPAATQG-KK**A**VLGK	0.47	**0.54**	0.65	*1.00*		0.60	**0.72**	0.91	*-1.8*
V29A	MNRGVPFRHLLLVLQLALLPAATQG-KKA**A**LGK	0.57	**0.63**	0.70	*+1.2*		0.86	**1.12**	1.49	*-1.1*
L30A	MNRGVPFRHLLLVLQLALLPAATQG-KKAV**A**GK	0.70	**0.84**	1.04	*+1.6*		0.87	**1.08**	1.37	*-1.2*
G31A	MNRGVPFRHLLLVLQLALLPAATQG-KKAVL**A**K	0.63	**0.72**	0.84	*+1.3*		2.01	**2.56**	3.32	*+2.0*
K32A	MNRGVPFRHLLLVLQLALLPAATQG-KKAVLG**A**	1.01	**1.42**	1.69	*+2.6*		6.02	**17.6**	>50	*+14*

HEK293T cells were transiently transfected with (mutant) constructs of FL huCD4 (left) or huCD4$_{1-32}$-YFP (right), and incubated with a serial 1:5 dilution of CADA compound (ranging from 50 to 0.016 µM) for 48h. The N, h and C-region of the signal peptide are indicated. For the wild-type (WT) huCD4 sequence, residues of the hydrophobic h-region are shown in bold face and underlined, whereas residues of the mature protein are indicated in blue.

[†]Expression of huCD4 protein was measured with flow cytometry by using a fluorescently labeled anti-CD4 antibody.

[‡]The IC_{50} (given in µM) corresponds to the CADA concentration that reduces the huCD4 signal to 50% of the untreated control.

[§]Mean IC_{50} and 95% confidence interval values (low/high) are interpolated from a four parameter dose-response curve, fitted to data from at least three replicate experiments.

[¶]FC: fold change (relative to WT control). The ratio of the mean IC_{50} value of each mutant over WT. "+" indicates an increase in IC_{50} value (less sensitive to CADA), and "-" a decrease (more sensitive to CADA).
[#]Same as in ([†]) but for the secretion of huCD4$_{1-32}$-YFP. Protein expression was determined by measuring the YFP fluorescence with flow cytometry.

Figure 1

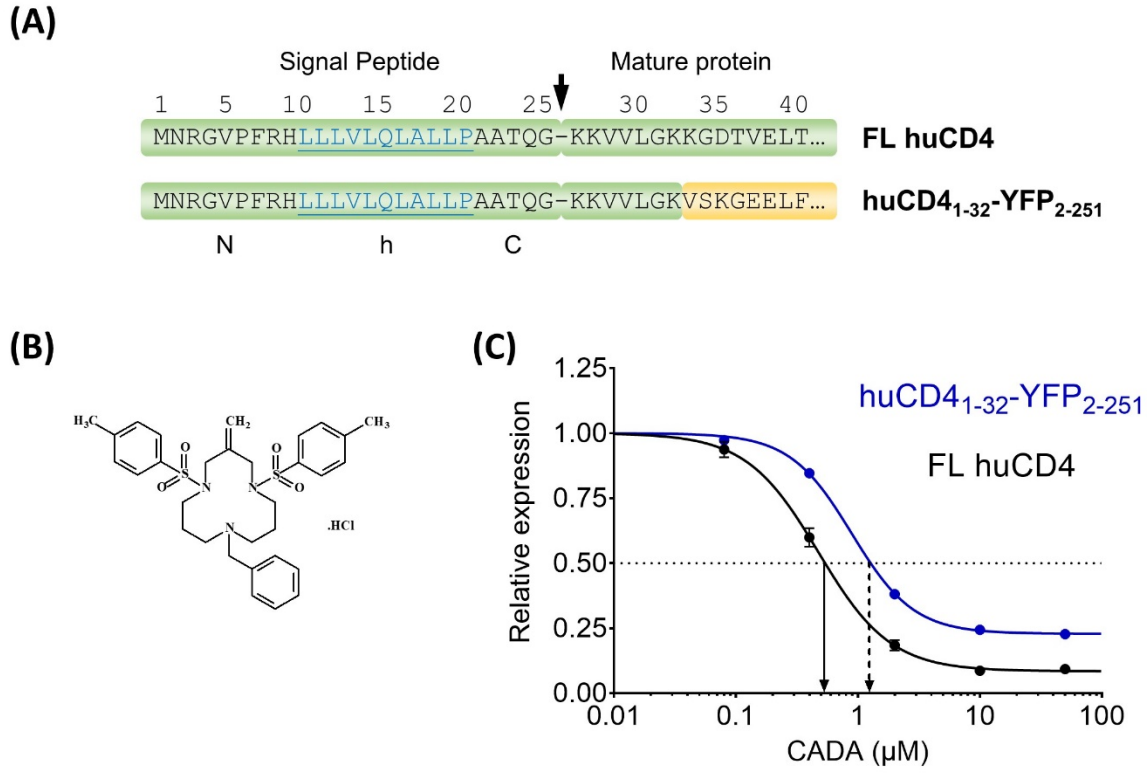

(A)
Signal Peptide
Mature protein
1 5 10 15 20 25 30 35 40
MNRGVPFRHLLLVLQLALLPAATQG-KKVVLGKKGDTVELT…
FL huCD4
MNRGVPFRHLLLVLQLALLPAATQG-KKVVLGKVSKGEELF…
huCD4$_{1-32}$-YFP$_{2-251}$
N
h
C
(B)
(C)
huCD4$_{1-32}$-YFP$_{2-251}$
FL huCD4
Relative expression
1.25
1.00
0.75
0.50
0.25
0.00
0.01
0.1
1
10
100
CADA (μM)

Figure 2

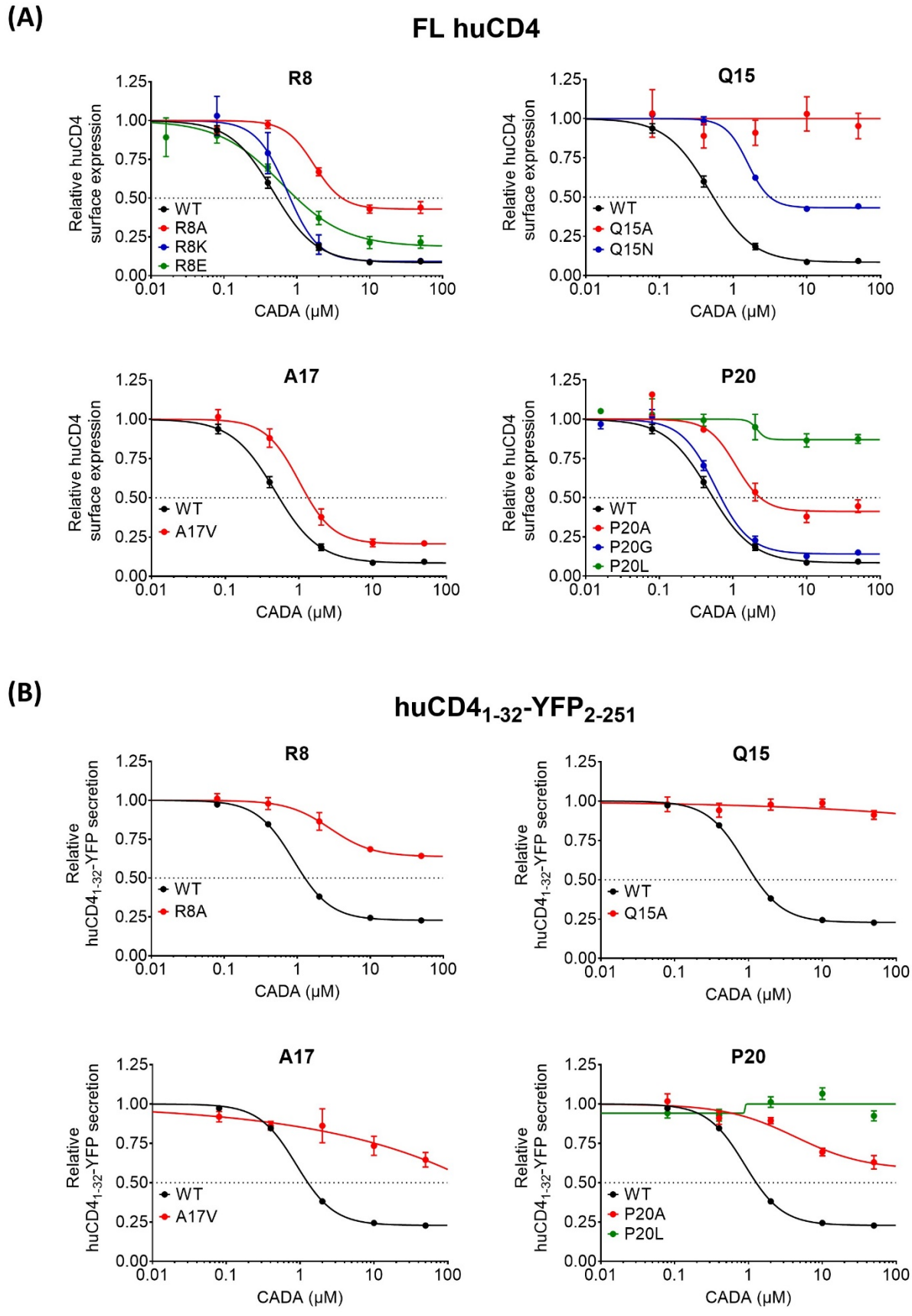

(A)
FL huCD4
R8
Relative huCD4 surface expression
WT
R8A
R8K
R8E
CADA (μM)
Q15
WT
Q15A
Q15N
A17
WT
A17V
P20
WT
P20A
P20G
P20L
(B)
huCD4$_{1-32}$-YFP$_{2-251}$
Relative huCD4$_{1-32}$-YFP secretion
R8
WT
R8A
Q15
WT
Q15A
A17
WT
A17V
P20
WT
P20A
P20L

Figure 3

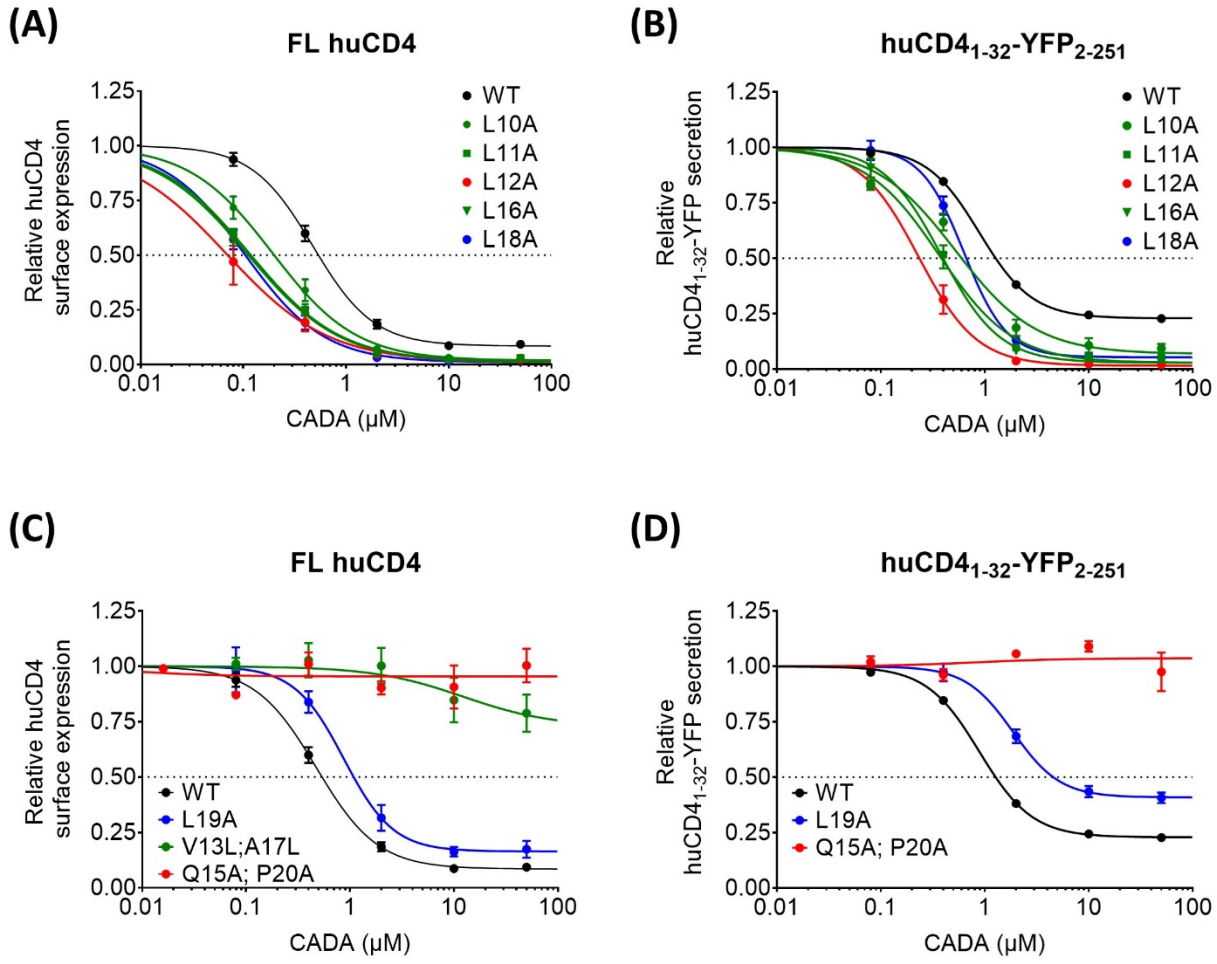

(A)
FL huCD4
Relative huCD4 surface expression
CADA (μM)
WT
L10A
L11A
L12A
L16A
L18A
(B)
huCD4$_{1-32}$-YFP$_{2-251}$
Relative huCD4$_{1-32}$-YFP secretion
CADA (μM)
WT
L10A
L11A
L12A
L16A
L18A
(C)
FL huCD4
Relative huCD4 surface expression
CADA (μM)
WT
L19A
V13L;A17L
Q15A; P20A
(D)
huCD4$_{1-32}$-YFP$_{2-251}$
Relative huCD4$_{1-32}$-YFP secretion
CADA (μM)
WT
L19A
Q15A; P20A

Figure 4

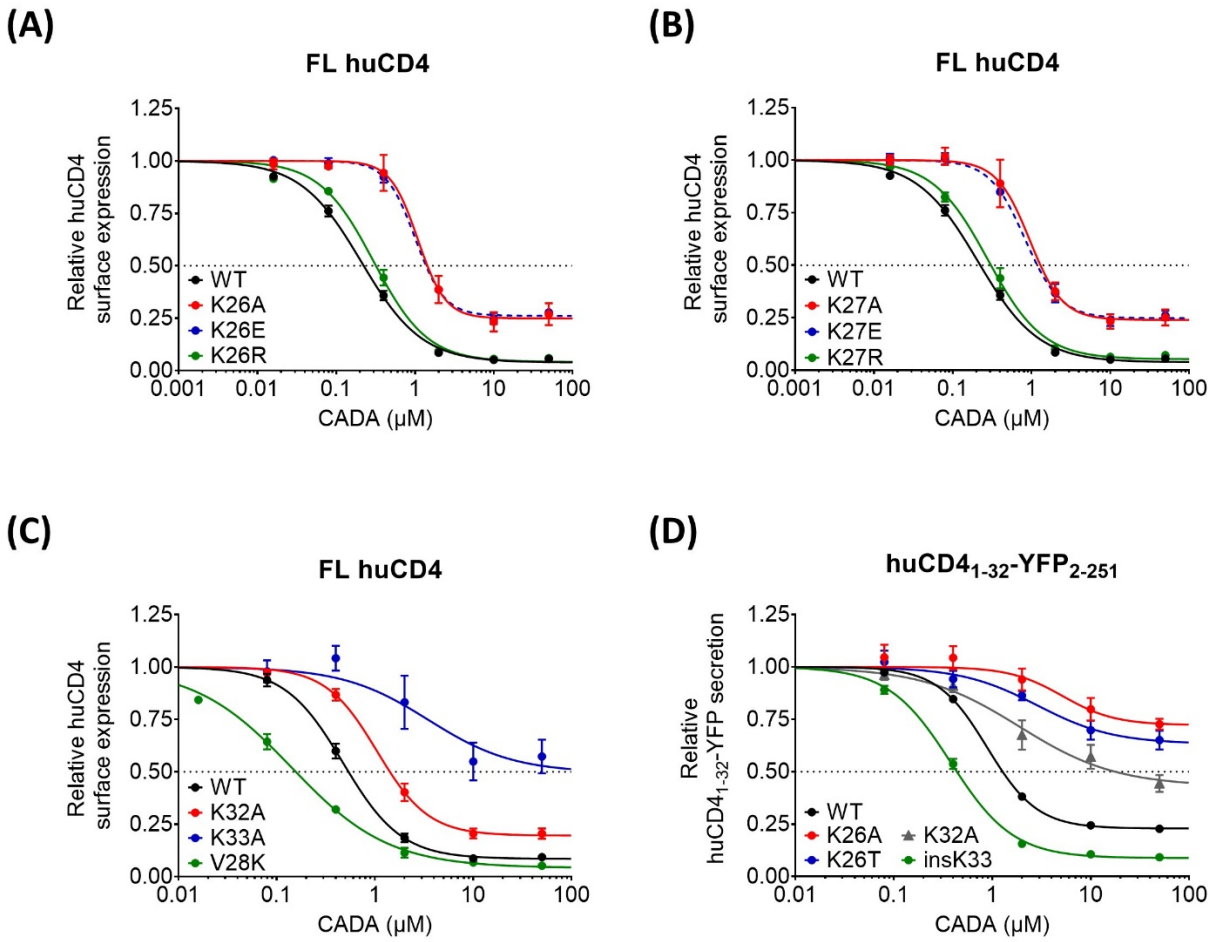

(A)
FL huCD4
Relative huCD4 surface expression
1.25
1.00
0.75
0.50
0.25
0.00
0.001
0.01
0.1
1
10
100
CADA (μM)
WT
K26A
K26E
K26R
(B)
FL huCD4
Relative huCD4 surface expression
CADA (μM)
WT
K27A
K27E
K27R
(C)
FL huCD4
Relative huCD4 surface expression
CADA (μM)
WT
K32A
K33A
V28K
(D)
huCD4$_{1-32}$-YFP$_{2-251}$
Relative huCD4$_{1-32}$-YFP secretion
CADA (μM)
WT
K26A
K26T
K32A
insK33

Figure 5

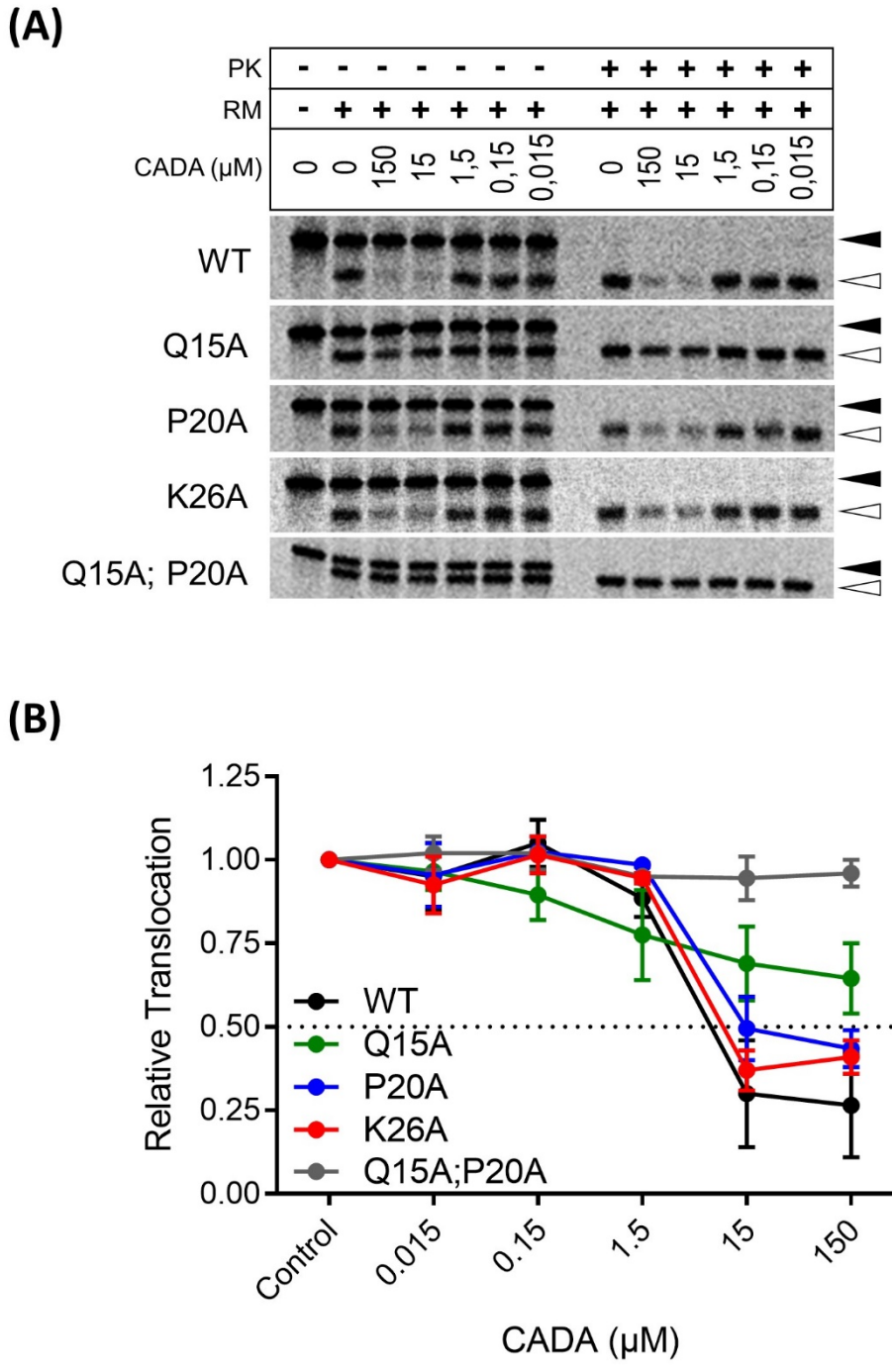

(A)
PK
RM
CADA (μM)
0
0
150
15
1,5
0,15
0,015
0
150
15
1,5
0,15
0,015
WT
Q15A
P20A
K26A
Q15A; P20A
(B)
Relative Translocation
1.25
1.00
0.75
0.50
0.25
0.00
WT
Q15A
P20A
K26A
Q15A;P20A
Control
0.015
0.15
1.5
15
150
CADA (μM)

Figure 6

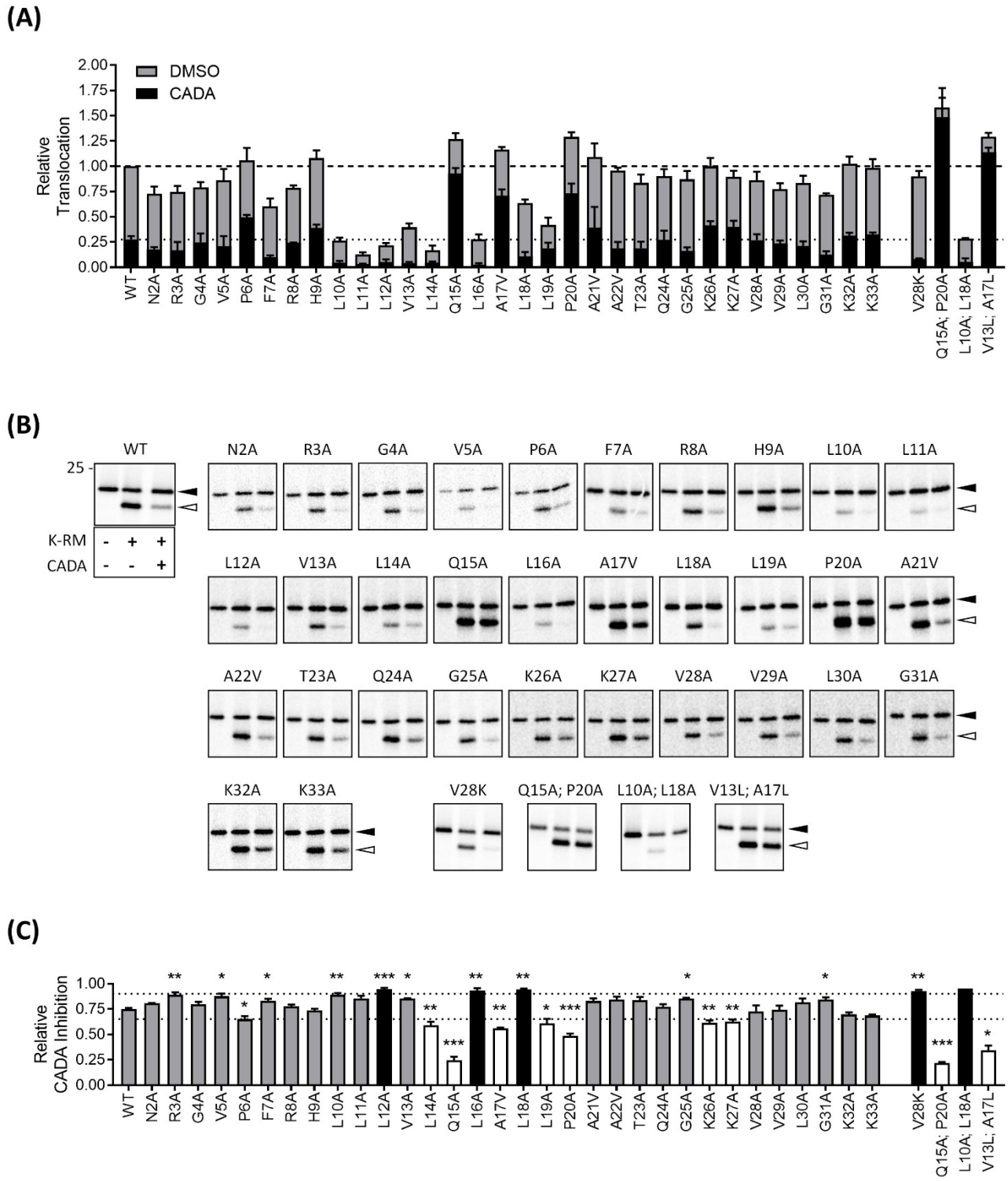

(A)
DMSO
CADA
Relative Translocation
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
WT N2A R3A G4A V5A P6A F7A R8A H9A L10A L11A L12A V13A L14A Q15A L16A A17V L18A L19A P20A A21V A22V T23A Q24A G25A K26A K27A V28A V29A L30A G31A K32A K33A V28K Q15A; P20A L10A; L18A V13L; A17L
(B)
WT N2A R3A G4A V5A P6A F7A R8A H9A L10A L11A
25
K-RM - + +
CADA - - +
L12A V13A L14A Q15A L16A A17V L18A L19A P20A A21V
A22V T23A Q24A G25A K26A K27A V28A V29A L30A G31A
K32A K33A V28K Q15A; P20A L10A; L18A V13L; A17L
(C)
Relative CADA Inhibition
1.00
0.75
0.50
0.25
0.00
WT N2A R3A G4A V5A P6A F7A R8A H9A L10A L11A L12A V13A L14A Q15A L16A A17V L18A L19A P20A A21V A22V T23A Q24A G25A K26A K27A V28A V29A L30A G31A K32A K33A V28K Q15A; P20A L10A; L18A V13L; A17L

Figure7

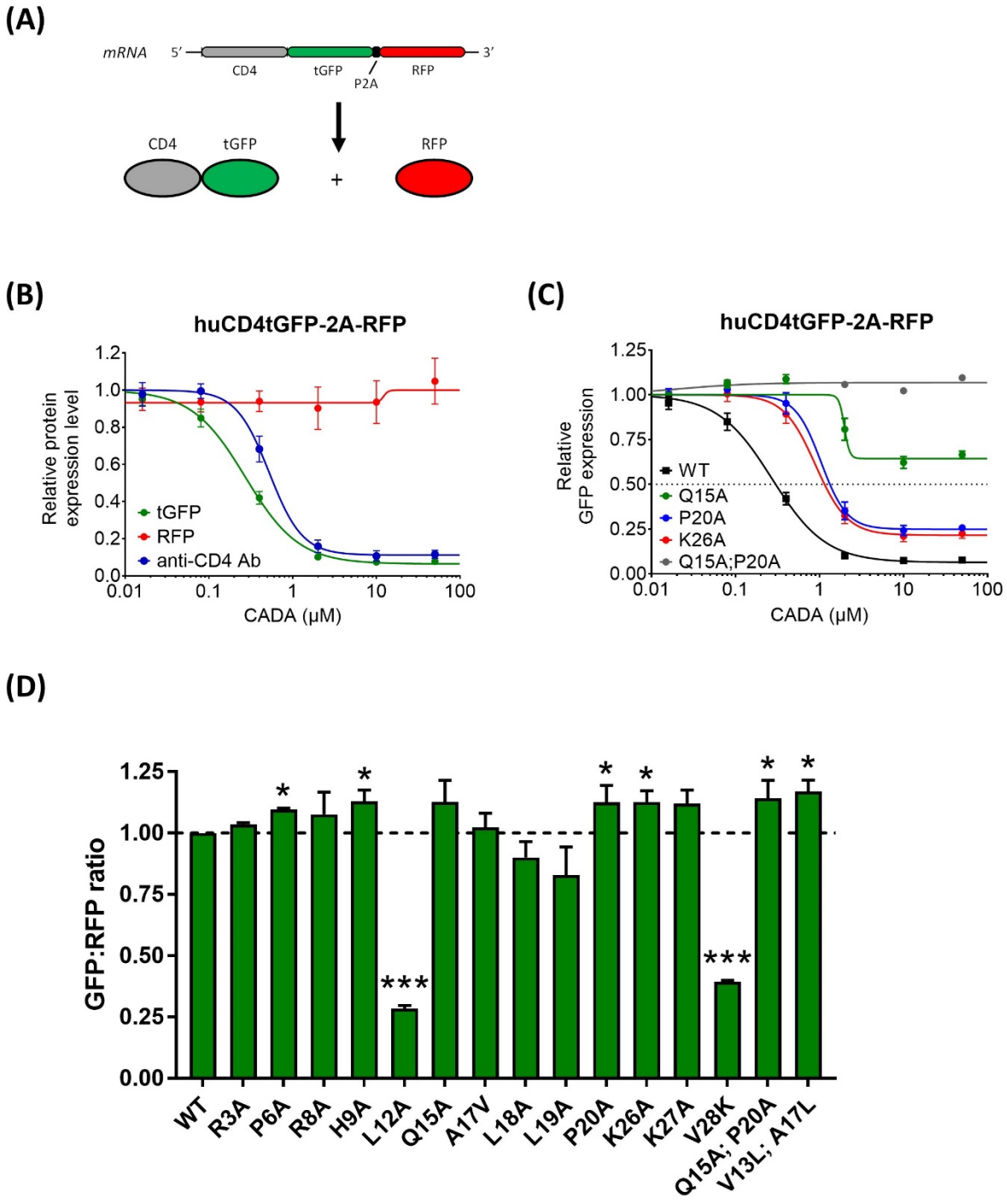

(A)
mRNA
5′
3′
CD4
tGFP
P2A
RFP
CD4
tGFP
+
RFP
(B)
huCD4tGFP-2A-RFP
Relative protein expression level
tGFP
RFP
anti-CD4 Ab
CADA (μM)
(C)
huCD4tGFP-2A-RFP
Relative GFP expression
WT
Q15A
P20A
K26A
Q15A;P20A
CADA (μM)
(D)
GFP:RFP ratio
WT
R3A
P6A
R8A
H9A
L12A
Q15A
A17V
L18A
L19A
P20A
K26A
K27A
V28K
Q15A; P20A
V13L; A17L

Figure Legends

Figure 1. Constructs used in the alanine scan mutagenesis. A, Amino acid sequence of the N-terminal part of wild-type full length (FL) huCD4 (top) and huCD4$_{1-32}$-YFP fusion protein (bottom), with 25 residues of the signal peptide and 15 residues of the mature protein. The arrow head indicates the signal peptide cleavage site. Residues from huCD4 are boxed in green, whereas those from YFP are boxed in yellow. The N-region, h-region and C-region are indicated below the signal peptide sequence. The residues of the hydrophobic h-region are in blue and underlined. B, Chemical structure of cyclotriazadisulfonamide (CADA). C, Dose-response curves of CADA for the wild-type constructs from (A). HEK293T cells were transiently transfected with construct plasmids and incubated with CADA for 48h. Tested CADA concentrations were 50, 10, 2, 0.4 and 0.08 µM. Expression of cellular huCD4 and secretion of huCD4$_{1-32}$-YFP was then measured with flow cytometry, and normalized to the DMSO control. A four parameter dose-response curve was fitted to data from at least three replicate experiments. Values are mean ± SD; n ≥ 3. The arrows indicate the mean IC_{50} value for both constructs (0.54 µM for FL huCD4 and 1.27 µM for huCD4$_{1-32}$-YFP).

Figure 2. huCD4 SP mutants differentially respond to CADA in transfected cells. A, Four parameter dose-response curves of different SP mutants (at position R8, Q15, A17 and P20) of the FL huCD4 construct. HEK293T cells were transiently transfected with the mutant constructs and incubated with different CADA concentrations for 48 hours as explained in the legend to figure 1C. B, Same as in (A) but for different alanine scan mutants of the huCD4$_{1-32}$-YFP construct.

Figure 3. Alteration in hydrophobicity of the huCD4 SP affects CADA sensitivity. Four parameter dose-response curves of different SP mutants of the h-region in FL huCD4 (A, C) and in huCD4$_{1-32}$-YFP (B, D). HEK293T cells were transiently transfected with the mutant constructs and incubated with different CADA concentrations for 48 hours as explained in the legend to figure 1C. A-B, Different leucine mutants of the SP with enhanced CADA susceptibility, with the most sensitive leucine mutant L12A indicated in red. C-D, Single and double mutants of the SP with reduced CADA sensitivity.

Figure 4. Contribution of the charged residues in the N-terminus of the mature protein domain towards CADA sensitivity. Four parameter dose-response curves of different lysine mutants of the mature protein in FL huCD4 (A-C) and in huCD4$_{1-32}$-YFP (D). HEK293T cells were transiently transfected with the mutant constructs and incubated with different CADA concentrations for 48 hours as explained in the legend to Figure 1C.

Figure 5. Co-translational translocation of different huCD4 mutants affected by CADA. A, Radiolabeled cell-free *in vitro* translation of truncated huCD4 D1D2 WT and SP mutants, treated with increasing doses of CADA. In the presence of rough microsomes (RM), the pre-protein is translocated and

protected from proteinase K (PK), and the signal peptide is cleaved from the mature protein chain (smaller apparent molecular weight). *Filled arrowhead*, unprocessed nascent chains (pre-protein with a predicted molecular mass of 22.5 kDa); *open arrowhead*, mature protein with cleaved SP (predicted molecular mass of 20 kDa). B, Quantified radioblots from (A). The relative translocation is given in arbitrary units relative to the untreated matching control sample which is set as 1.0. Values are mean ± SD, n≥2.

Figure 6: Impact of alanine substitutions in the huCD4 SP on the cell-free *in vitro* translocation efficiency of CD4 mutants. A, Radiolabeled cell-free *in vitro* translation of truncated huCD4 D1D2 WT and SP mutants, treated with 15 µM of CADA or DMSO control. Translocation of control (DMSO, grey bars) and CADA-treated samples (black bars) normalized to untreated WT huCD4. The translocation efficiency of WT huCD4D1D2 is set at 1.00 and is indicated with a dotted line. Bars show mean ± SEM of two to five independent experiments. B, Representative radioblots from (A). *Filled arrowhead*, unprocessed nascent chains (pre-protein with a predicted molecular mass of 22.5 kDa); *open arrowhead*, mature protein with cleaved SP (predicted molecular mass of 20 kDa). K-RM: salt-washed rough microsomes. The molecular weight marker of 25 kDa is shown on the WT sample. Note that we can not exclude the possibility that some of the uncleaved species (filled arrowhead) might be translocated chains but not signal peptide cleaved. C, Inhibitory effect of CADA on the translocation of WT and mutant CD4, calculated on the samples from (A). For each mutant, the intensity of the SP-cleaved protein band of the CADA sample in the radioblot was normalized to its corresponding DMSO control (*open arrowhead* in radioblots) to calculate the relative inhibition (complete inhibition is set as 1.0). Dotted lines are arbitrary set at 0.65 and 0.90 to visualize the window of variation for WT. Mutants for which the CADA inhibition was between these values are represented in grey. Bars with inhibition > 0.90 are indicated in black (i.e., mutants highly sensitive to CADA), bars with inhibition < 0.65 are shown in white (i.e., resistant mutants). Bars show mean ± SEM of two to five independent experiments. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, compared to WT control (two-tailed unpaired t test with Welch's correction).

Figure 7: Impact of alanine substitutions in the SP on the cellular expression efficiency of huCD4 tGFP-2A-RFP mutants**.** A, Schematic representation of the expected mRNA and protein products of the huCD4tGFP-2A-RFP construct. B, Four parameter dose-response curves for CADA of WT huCD4tGFP-2A-RFP. Levels of CD4 were simultaneously determined by GFP measurement (green line) and by staining of surface huCD4 with an APC-labeled anti-CD4 antibody (blue line). In parallel, the RFP level in each cell was also quantified by flow cytometry (red line). Protein levels in CADA-treated samples are normalized to the DMSO control (set at 1.0). Curves are fitted to data from four replicate experiments. Values are mean ± SD. C, Four parameter dose-response curves for CADA on tGFP

expression of WT and different mutant huCD4tGFP-2A-RFP [same as in (B)]. Curves are fitted to data from at least three replicate experiments. Values are mean ± SD. D, GFP:RFP ratio of alanine mutants in the absence of CADA, normalized to WT control (set as 1.0). The expression level of RFP of each mutant was used to compare with WT. Values below 1.00 indicate reduced expression level of untreated huCD4-GFP as compared to the WT construct. Bars are mean ± SD; n ≥ 3. *P < 0.05, ***P < 0.001, compared to WT control (two-tailed unpaired t test with Welch's correction).

Figure S1: Impact of alanine substitutions in the SP on the cellular expression efficiency of huCD4 tGFP-2A-RFP mutants. Four parameter dose-response curves for CADA on huCD4 expression of WT and different mutant huCD4tGFP-2A-RFP by staining of surface huCD4 with an APC-labeled anti-CD4 antibody (same samples as in Figure 7C). Curves are fitted to data from at least three replicate experiments. Values are mean ± SD.

Supplemental materials

Table S1. CADA sensitivity of additional single and double mutants of huCD4.

Mutant	1 5 10 15 20 25 32	Cellular huCD4 IC_{50} (µM) Low	**Mean**	high
R8K	MNRGVPF**K**HLLLVLQLALLPAATQG-KKVVLGKK	0.52	**0.81**	1.46
R8E	MNRGVPF**E**HLLLVLQLALLPAATQG-KKVVLGKK	0.68	**0.99**	1.54
L10A;L18A	MNRGVPFRH**A**LLVLQLA**A**LPAATQG-KKVVLGKK	0.04	**0.06**	0.08
L16A;L18A	MNRGVPFRHLLLVLQ**A**A**A**LPAATQG-KKVVLGKK	0.09	**0.11**	0.15
V13L;A17L	MNRGVPFRHLLL**L**LQL**L**LLPAATQG-KKVVLGKK	>50	**>50**	>50
Q15N	MNRGVPFRHLLLVL**N**LALLPAATQG-KKVVLGKK	2.46	**3.17**	4.85
Q15A;P20A	MNRGVPFRHLLLVL**A**LALL**A**AATQG-KKVVLGKK	>50	**>50**	>50
P20G	MNRGVPFRHLLLVLQLALL**G**AATQG-KKVVLGKK	0.58	**0.69**	0.87
P20L	MNRGVPFRHLLLVLQLALL**L**AATQG-KKVVLGKK	>50	**>50**	>50
V28K	MNRGVPFRHLLLVLQLALLPAATQG-KK**K**VLGKK	0.13	**0.16**	0.19
K33A	MNRGVPFRHLLLVLQLALLPAATQG-KKKVLGK**A**	>50	**>50**	>50

HEK293T cells were transiently transfected with (mutant) constructs and incubated with a serial 1:5 dilution of CADA compound (ranging from 50 to 0.016 µM) for 48h. Expression of huCD4 protein was measured with flow cytometry by using a fluorescently labeled anti-CD4 antibody. The IC_{50} (given in µM) corresponds to the CADA concentration that reduces the huCD4 signal to 50% of the untreated control. Mean IC_{50} and 95% confidence interval values (low/high) are interpolated from a four parameter dose-response curve, fitted to data from at least three replicate experiments.

Table S2. Differential CADA sensitivity of alanine scan mutants of huCD4tGFP-2A-RFP.

Mutant	1 5 10 15 20 25 32	Cellular GFP IC_{50} (µM) Low	**Mean**	high
WT	MNRGVPFRHLLLVLQLALLPAATQG-KKVVLGK	0.28	**0.30**	0.32
R3A	MN**A**GVPFRHLLLVLQLALLPAATQG-KKVVLGK	0.22	**0.25**	0.28
P6A	MNRGV**A**FRHLLLVLQLALLPAATQG-KKVVLGK	0.91	**1.22**	1.51
R8A	MNRGVPF**A**HLLLVLQLALLPAATQG-KKVVLGK	0.68	**0.84**	1.10
H9A	MNRGVPFR**A**LLLVLQLALLPAATQG-KKVVLGK	0.70	**0.78**	0.88
L12A	MNRGVPFRHLL**A**VLQLALLPAATQG-KKVVLGK	0.06	**0.07**	0.08
Q15A	MNRGVPFRHLLLVL**A**LALLPAATQG-KKVVLGK	>50	**>50**	>50
A17V	MNRGVPFRHLLLVLQL**V**LLPAATQG-KKVVLGK	0.55	**0.67**	0.91
L18A	MNRGVPFRHLLLVLQLA**A**LPAATQG-KKVVLGK	0.06	**0.07**	0.08
L19A	MNRGVPFRHLLLVLQLAL**A**PAATQG-KKVVLGK	0.30	**0.35**	0.41
P20A	MNRGVPFRHLLLVLQLALL**A**AATQG-KKVVLGK	1.14	**1.33**	1.52
K26A	MNRGVPFRHLLLVLQLALLPAATQG-**A**KVVLGK	1.01	**1.15**	1.31
K27A	MNRGVPFRHLLLVLQLALLPAATQG-K**A**VVLGK	0.98	**1.11**	1.27
V28K	MNRGVPFRHLLLVLQLALLPAATQG-KK**K**VLGK	0.05	**0.06**	0.07
Q15A;P20A	MNRGVPFRHLLLVL**A**LALL**A**AATQG-KKVVLGK	>50	**>50**	>50
V13L;A17L	MNRGVPFRHLLL**L**LQL**L**LLPAATQG-KKVVLGK	1.98	**2.61**	4.18

HEK293T cells were transiently transfected with (mutant) constructs and incubated with a serial 1:5 dilution of CADA compound (ranging from 50 to 0.016 µM) for 24h. Expression of huCD4 protein was determined with flow cytometry by measuring the GFP fluorescence. The IC_{50} (given in µM) corresponds to the CADA concentration that reduces the GFP signal to 50% of the untreated control. Mean IC_{50} and 95% confidence interval values (low/high) are interpolated from a four parameter dose-response curve, fitted to data from at least three replicate experiments.

Figure S1

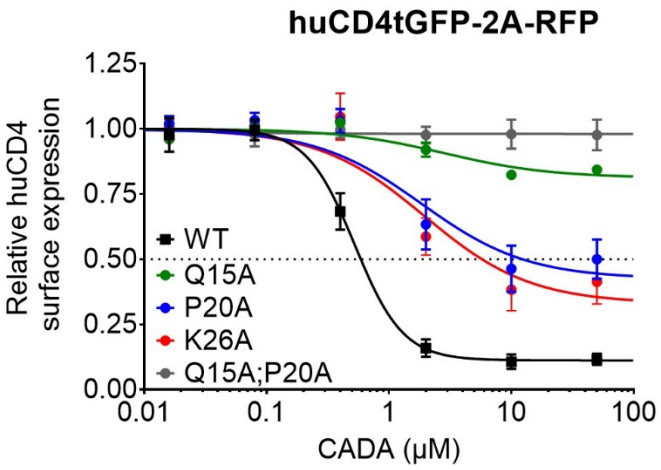


Figure S1: Impact of alanine substitutions in the SP on the cellular expression efficiency of huCD4 tGFP-2A-RFP mutants. Four parameter dose-response curves for CADA on huCD4 expression of WT and different mutant huCD4tGFP-2A-RFP by staining of surface huCD4 with an APC-labeled anti-CD4 antibody (same samples as in Figure 7C). Curves are fitted to data from at least three replicate experiments. Values are mean ± SD.